AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 108:1–39 (1999)

Mitochondrial DNA Variation in Koryaks and Itel'men: Population Replacement in the Okhotsk Sea–Bering Sea Region During the Neolithic

THEODORE G. SCHURR,[1,2] REM I. SUKERNIK,[3]
YELENA B. STARIKOVSKAYA,[3] AND DOUGLAS C. WALLACE[1,2]*
[1]*Department of Anthropology, Emory University, Atlanta, Georgia 30322*
[2]*Center for Molecular Medicine, Emory University, Atlanta, Georgia 30322*
[3]*Laboratory of Human Molecular Genetics, Institute of Cytology and Genetics, Siberian Division, Russian Academy of Sciences, Novosibirsk 630090, Russia*

KEY WORDS haplotype; haplogroup; control region; sublineage; Kamchatka

ABSTRACT In this study, we analyzed the mitochondrial DNA (mtDNA) variation in 202 individuals representing one Itel'men and three Koryak populations from different parts of the Kamchatka peninsula. All mtDNAs were subjected to high resolution restriction (RFLP) analysis and control region (CR) sequencing, and the resulting data were combined with those available for other Siberian and east Asian populations and subjected to statistical and phylogenetic analysis. Together, the Koryaks and Itel'men were found to have mtDNAs belonging to three (A, C, and D) of the four major haplotype groups (haplogroups) observed in Siberian and Native American populations (A–D). In addition, they exhibited mtDNAs belonging to haplogroups G, Y, and Z, which were formerly called "Other" mtDNAs. While Kamchatka harbored the highest frequencies of haplogroup G mtDNAs, which were widely distributed in eastern Siberian and adjacent east Asian populations, the distribution of haplogroup Y was restricted within a relatively small area and pointed to the lower Amur River–Sakhalin Island region as its place of origin. In contrast, the pattern of distribution and the origin of haplogroup Z mtDNAs remained unclear. Furthermore, phylogenetic and statistical analyses showed that Koryaks and Itel'men had stronger genetic affinities with eastern Siberian/east Asian populations than to those of the north Pacific Rim. These results were consistent with colonization events associated with the relatively recent immigration to Kamchatka of new tribes from the Siberian mainland region, although remnants of ancient Beringian populations were still evident in the Koryak and Itel'men gene pools. Am J Phys Anthropol 108:1–39, 1999.

Grant sponsor: Wenner-Gren Foundation for Anthropological Research; Grant numbers: GR 5745, GR 5680, GR 5680-S, GR 6077; Grant sponsor: The National Science Foundation; Grant number: 9414900; Grant sponsor: The Leakey Foundation; Grant sponsor: Sigma Xi; Grant sponsor: INTAS; Grant number: 96–1766; Grant sponsor: Russian Fund for Basic Research; Grant number: 97–04–49816; Grant sponsor: J. Worley Brown Fellowship Fund; Grant sponsor: National Institutes of Health; Grant numbers: GM49615, NS21328, HL45572, AG10130, AG13154; Grant sponsor: Emory Clinical Research Center; Grant number: BR00039.

*Correspondence to: Douglas C. Wallace, Center for Molecular Medicine, Emory University School of Medicine, 1462 Clifton Road N.E., Atlanta, GA 30322.
E-mail: dwallace@gmm.gen.emory.edu

Received 20 May 1997; accepted 3 October 1998.

In the past 10 years, we have conducted a number of studies in which the mitochondrial DNA (mtDNA) variation of aboriginal populations in Siberia and the Americas was used to trace their origins and affinities (Wallace et al., 1985; Schurr et al., 1990; Torroni et al., 1992, 1993a,b, 1994a,b; Sukernik et al., 1996; Starikovskaya et al., 1998). These analyses showed that the mtDNAs observed in contemporary New World populations were a subset of Asian haplotypes which consisted of primarily four mtDNA lineages, or haplogroups, designated A, B, C, and D. Within these mtDNA lineages, only a small number of haplotypes were found to be shared between Asian and Native American populations, suggesting that a limited number of founders gave rise to ancestral Native American populations. The sequence divergence values for the haplogroups present in Native Americans further indicated that ancestral populations bringing at least haplogroups A, C, and D arrived early in the New World, around 35,000–25,000 years before present (YBP), and that haplogroup B might represent a second, later migration which contributed mtDNAs to the genetic stock of Amerindians (Torroni et al., 1992, 1993a, 1994a).

Other studies of mtDNA variation in Asian and Native American populations are at variance with some of these interpretations. Ward et al. (1991) and Horai et al. (1993) argued that extensive bottlenecks had not caused limited mtDNA variation among Native American groups since they observed control region (CR) sequence diversity within Native American tribes that was similar to levels found in Asian populations. They also detected four major clusters of CR sequences equivalent to haplogroups A–D, although Horai et al. (1993) claimed that each haplogroup represented a separate migration into the New World which came between 21,000 and 14,000 YBP. On the other hand, Shields et al. (1993) proposed a late entry time (16,000–14,000 YBP) of ancestral Amerindians into the New World, along with the later expansion of northern populations in the circumarctic region. Moreover, since haplogroups A–D are present in most Native American groups, it was argued that all of these mtDNA lineages were brought to the Americas in a single migratory wave (Bailliet et al., 1994; Merriwether et al., 1994, 1995; Kolman et al., 1996).

When the mtDNA variation in native Siberian populations was initially surveyed, only three of the four Asian mtDNA haplogroups (A, C, and D) were detected in these groups (Torroni et al., 1993a). With only a few exceptions (Petrishchev et al., 1993; Sukernik et al., 1996; Derenko and Shields, 1998), Siberian groups lacked haplogroup B mtDNAs. Central-east Asian populations, by contrast, had haplogroup B mtDNAs at polymorphic frequencies (Horai and Matsunaga, 1986; Ballinger et al., 1992; Harihara et al., 1992; Kolman et al., 1996; Merriwether et al., 1996), suggesting that this mtDNA lineage arose in this area of Asia. The mtDNA data also revealed the presence of additional haplogroups in northern Asia in the form of "Other" (non–haplogroup A–D) haplotypes, which appeared in all native Siberian groups except for the Siberian Eskimos (Torroni et al., 1993b; Sukernik et al., 1996). Some of these "Other" haplogroups appeared to be related to previously observed east Asian mtDNAs (Horai et al., 1984; Horai and Matsunaga, 1986; Ballinger et al., 1992; Torroni et al., 1993b), whereas the lineal associations of the remainder were unclear. Similarly, "Other" mtDNAs have been observed in Native American populations (Bailliet et al., 1994; Merriwether et al., 1995; Easton et al., 1996; Lorenz and Smith, 1996). However, it is unclear whether or not the Native American and Siberian "Other" mtDNAs are related to each other.

Because of the Kamchatka peninsula's geographic proximity to the Bering Strait region, the Aleutian Islands, and the Kurile Islands and northern Japan, the aboriginal populations inhabiting that peninsula may be important for clarifying northern Asian prehistory and the colonization of the New World. It was also possible that these populations were genetically influenced by the hypothesized migration which brought haplogroup B mtDNAs to the Americas (Torroni et al., 1992, 1993a,b). In addition, an analysis of Koryak and Itel'men mtDNA variation might also delineate the origins and affini-

ties of the "Other" haplotypes which had previously been observed in eastern Siberian and Native American populations (Torroni et al., 1993b).

However, aside from the classical anthropological surveys of aboriginal populations inhabiting eastern Siberia (Debets, 1951; Levin, 1958), few studies have attempted to delineate the biological variation of Kamchatkan groups. While some analyses of mtDNA variation in northeast Siberian groups have provided little conclusive evidence for population affinities in this region (Malyarchuk et al., 1994; Derenko and Shields, 1998), our initial analysis of Siberian populations revealed the presence of haplogroups A, C, and D in the Chukchi and "Koryaks" (Torroni et al., 1993b), implying a close linkage between them. However, no additional associations could be made because the data were obtained through partial haplotype analysis. In addition, few if any genetic studies of the Itel'men have been conducted, possibly due to their having significant levels of nonnative admixture with Russians.

To further elucidate the genetic affinities of eastern Siberian populations and their role in the peopling of the New World, we conducted a detailed molecular analysis of the mtDNA variation in Koryaks and Itel'men groups from the Kamchatka peninsula. The data obtained through high resolution restriction fragment length polymorphism (RFLP) analyses and CR sequencing were combined with similar data sets from the Chukchi and Siberian Eskimos of Chukotka, the Nivkhs and Udegeys of the lower Amur–northern Sakhalin region, and the Evenks of interior Siberia and subjected to statistical and phylogenetic analyses. Our results genetically reflect the recent immigrations to Kamchatka of ancestral Paleoasiatic populations from the Siberian mainland and their nearly complete replacement of the ancient Beringian populations which formerly inhabited this region.

POPULATIONS

The Koryaks

As noted by Stepan Krasheninnkov (1972: 193–195) in *Opisanie Zemli Kamchatki* (*Description of the Land of Kamchatka* [St. Petersburg, 1754]), the Koryaks were "divided into two nations, one called the reindeer Koryak, the other the settled Koryaks." The sedentary (Maritime) Koryaks established permanent settlements along rivers flowing into the Sea of Okhotsk and the Bering Sea and subsisted to varying degrees on hunting small sea mammals, fishing, and gathering plant and animal species from the littoral zone (Jochelson, 1908; Antropova, 1964a; Krashenninnkov, 1972). Their population was traditionally subdivided into eight territorial and dialectical subgroups whose members spoke different dialects of Paleoasiatic, or Chukotko-Kamchatkan, languages (Skorik 1965; Krauss, 1988).

At the turn of the seventeenth century, Maritime Koryaks occupied northern Kamchatka and the northeastern Okhotsk Sea region (Jochelson, 1908; Antropova, 1964a). However, the expansion of the Evens along the Okhotsk Sea coast in the seventeenth and eighteenth centuries gradually reduced the territory occupied by Maritime Koryaks (Levin and Vasiliev, 1964) and forced those inhabiting this region to shift to reindeer breeding—hence a nomadic way of life (Jochelson, 1908; Levin and Vasiliev, 1964; Arutiunov, 1988). By the turn of the twentieth century, these Reindeer Koryaks inhabited the forest tundra zone of northwestern Kamchatka and the Penzhina River basin and the mountain tundras in the northeastern part of the Kamchatka mainland (Jochelson, 1908; Antropova, 1964a). Interestingly, Reindeer Koryaks spoke a language that was nearly unintelligible to Maritime Koryaks, while Reindeer Koryak was close enough to Chukchi in vocabulary and morphology that it was mutually intelligible to native Chukchi speakers (Antropova, 1964a; Krasheninnikov, 1972; Vdovin, 1973; Arutiunov, 1988; Krauss, 1988).

Today, the majority of Koryaks reside within the borders of the Koryak Autonomous Region, which lies between 56° and 65°N and 158° and 174°E in northeastern Siberia (Fig. 1). According to the 1989 All-Union census, the total number of Koryaks in the Koryak Autonomous Region was 6,572, with approximately half of these being Reindeer Koryaks, whereas a century ago their population numbered 7,284, of whom 2,913

were Reindeer Koryaks (Krushanov, 1993). These figures indicate that the Koryaks have largely maintained their tribal integrity since the Russian entry into northeast Siberia.

The Itel'men

During the initial period of Russian colonization, the Itel'men inhabited much of the Kamchatka peninsula. Their populations extended from its southern tip, where Ainu populations resided, to the north around the Uka and Tigil' Rivers, where they interspersed with Koryaks, with their main population being concentrated in the Kamchatka River basin (Antropova, 1964b; Krasheninnikov, 1972) (Fig. 1). Those groups living in the interior subsisted primarily through fishing along the rivers running through the peninsula, although those inhabiting the east coast also hunted sea mammals and those in the southern tip of the peninsula hunted whales (Antropova, 1964b; Krasheninnikov, 1972). Itel'men populations were also divided by language into territorial subgroups which resided in the western, central/east coast, and southern parts of the peninsula (Antropova, 1964b; Arutiunov, 1988) (Fig. 1). Although sometimes considered part of the Paleoasiatic linguistic group, the Itel'men language is distinct from the Koryak language and may have become similar to it through population contact and linguistic borrowing (Krasheninnikov, 1972; Vdovin, 1973; Arutiunov, 1988; Krauss, 1988; Krushanov, 1990).

Since contact with Russians, Itel'men populations have declined precipitously, falling from approximately 12,000 individuals in the early eighteenth century to only 814 persons reported in the 1926–27 census (Antropova, 1964b). During this time, they became increasingly culturally and genetically assimilated with Russians such that only a small subdivision of their original population has survived into this century. Its members live primarily in the southwest corner of the Koryak Autonomous Region, where they have retained their native language and ethnic identity (Antropova, 1964b; Arutiunov, 1988).

In this context, it should be noted that the sedentary natives of southern Kamchatka were referred to as Kamchadals throughout the colonial period. Due to extensive admixture with ethnic Russians, the term *Kamchadal* was applied to both Itel'men and mixed Itel'men-Russian individuals as well as the descendants of Russian Cossacks and peasants who settled in Kamchatka in the eighteenth and nineteenth centuries (Antropova, 1964b; Vdovin 1973; Arutiunov, 1988; Murashko, 1994) and thus did not clearly discriminate between the indigenous and immigrant populations of this region. During the early Soviet period, national regionalization created new administrative borders which officially divided these Kamchadal descendants into two groups: the Itel'men, who lived within the boundaries of the Koryak Autonomous Region, and the Kamchadals, who lived south of the borders of the Koryak Autonomous Region, with members of the latter group being classified as "russified" and deprived of their "small indigenous people" status (Arutiunov, 1988; Murashko, 1994).

Sample collection

In July–August 1993, blood samples were collected from 104 Koryaks residing in three geographically proximate villages, Karaga, Ossora, and Tymlat, located in the Karaginskiy District of the Koryak Autonomous Region. Almost all of the Koryaks living in Ossora and Tymlat (57) represented the Aluitor subgroup. Of this total, 27 and 12 individuals were born in the villages of Rekinniki and Anapka, respectively, both of which were closed in the early 1960s. Once considered to speak a distinct language (Skorik, 1965), the Aluitor were the largest territorial group of Koryaks and occupied the whole Kamchatka isthmus and adjacent Bering Sea coast, combining small-scale reindeer herding with sea mammal hunting and fishing (Arutiunov, 1988).

Whereas approximately half of the Karaginskiy District sample consisted of Aluitor Koryaks, the remaining half represented the Karagin Koryaks. Speaking a slightly modified dialect of the Aluitor language (Vdovin, 1973), the Karagin Koryaks traditionally occupied the territory south of Tymlat, including Karaginskiy Island and settlements scattered along the Bering Sea coast

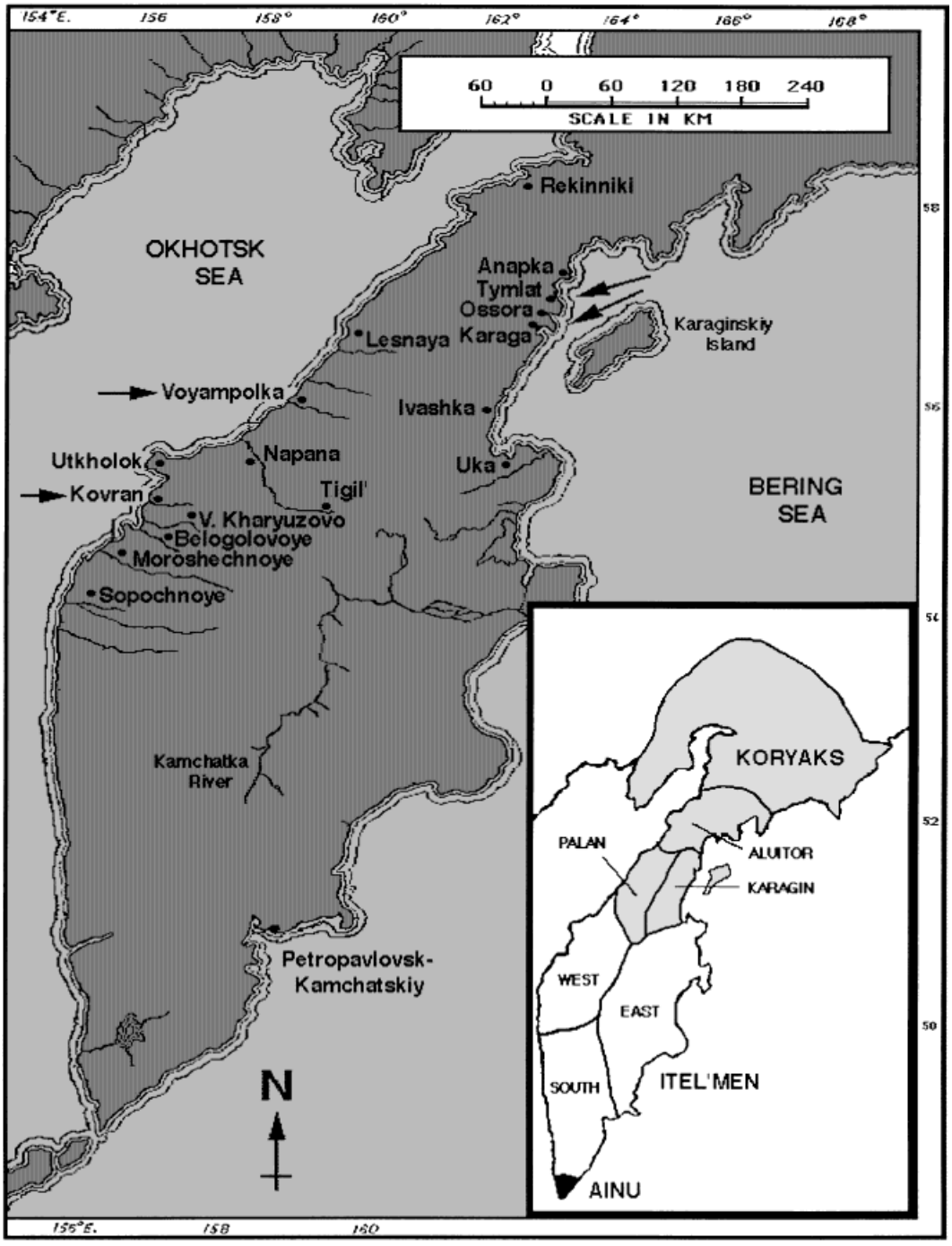

Fig. 1. A map of Kamchatka showing the locations of the villages in which fieldwork was conducted in 1993 and 1996 (with arrows) and other traditional settlements of Koryak and Itel'men, the majority of which no longer exist. **Inset:** The traditional territories of the Koryaks (grey) and Itel'men (white) around the beginning of the eighteenth century, with the geographic locations of the dialectic subgroups for each population indicated on the Kamchatka peninsula.

down to the village of Uka on the Ukinskaya Inlet. However, they now live in the villages of Ossora and Karaga. Although no longer in existence, some of these traditional villages (including Ivashka and Uka) were once occupied by Itel'men, who apparently merged with the Karagin Koryaks by the middle of the nineteenth century (Antropova, 1964a; Vdovin, 1973; Krushanov, 1993).

In June 1996, blood samples were obtained from 51 Koryak and 47 Itel'men residing in the villages of Voyampolka and Kovran, which are located in the Tigil'skiy District of the Koryak Autonomous Region. The Voyampolka sample was comprised of persons having mixed Maritime and Reindeer Koryak origins but who could be considered to belong to the Palan subgroup (Krashenninnkov, 1972; Vdovin, 1973; Krushanov, 1993). The Palan Koryaks traditionally lived in several settlements scattered along the Okhotsk Sea coast of the peninsula between Voyampolka in the south and Lesnaya in the north but now reside mostly in these two villages.

The Itel'men samples were obtained from individuals living in Kovran. All of these persons were born in or derived from one of a number of traditional settlements scattered along the Okhotsk Sea coast between Sopochnoye in the south and Tigil' in the north, including the villages of Kovran, Napaná, Utkholok, Moroshechnoye, Belogolovoye, and Verkhneye Kharyuzovo. With the exception of Kovran and Verkhneye Kharyuzovo, none of these former Itel'men villages exist today (Fig. 1).

Blood samples were collected from each participant with informed consent in two sets of 10 ml ACD anti-coagulant tubes and kept refrigerated in the local hospitals until shipped or brought back to Atlanta for molecular genetic analysis. All individuals who participated in these studies were interviewed about their family histories, which in turn were verified by senior members of the community for accuracy. Only those persons who lacked maternal and paternal Russian or nonrelated ancestry through three generations were selected for the collection of blood samples, although samples were also obtained from four Evens who were the marital partners of Koryak participants. Based on these genealogical data, we estimated that approximately half of the Koryaks and most of the Itel'men are of mixed Russian-Koryak or Russian-Itel'men ancestry, respectively, and consider themselves Koryak or Itel'men by nationality primarily because of their maternal ancestry.

METHODS

Blood sample processing

All blood samples were processed at Emory University at the Clinical Research Center. Two 10 ml tubes of blood from each individual were separated into their constituent cellular fractions through low-speed centrifugation. Lymphocytes were separated from the cellular fraction, and the residual platelet-rich plasmas from the centrifuged specimens were subsequently centrifuged in 15 ml Corning (Corning, NY) tubes at 5,000 rpm and 10°C for 20 min to precipitate the platelets, which were subsequently extracted for mtDNAs (Torroni et al., 1992).

High resolution restriction analysis

All Koryak and Itel'men DNA samples were subjected to high resolution restriction analysis. The entire mtDNAs of these samples were polymerase chain reaction (PCR) amplified in nine partially overlapping segments using standard oligonucleotide primers and PCR conditions (Torroni et al., 1992, 1993a). Each PCR segment was subsequently digested with 14 restriction enzymes (*AluI, AvaII, BamHI, DdeI, HaeII, HaeIII, HhaI, HincII, HinfI, HpaI, HpaII, MboI, RsaI,* and *TaqI*) to screen approximately 20% of the mtDNA sequence per individual. The resulting restriction fragments were resolved by electrophoresis in 1–2.5% NuSieve plus 1.0% SeaKem agarose (FMC BioProducts, Rockland, ME) gels and visualized by ethidium bromide staining. The restriction fragment length polymorphisms (RFLPs) detected by this analysis were mapped by the sequence comparison method (Johnson et al., 1983; Cann et al., 1984), with the combination of all RFLPs identified in a mtDNA defining its complete haplotype.

Control region sequencing

A total of 525 bp (np 16000–16525) encompassing CR hypervariable segment I (CR) was dideoxy-sequenced in 155 Koryak and 47 Itel'men mtDNAs. For each mtDNA, the CR was read in both directions by double-stranded sequencing of PCR products with an ABI 377A Automated DNA Sequencer (Applied Biosystems, Foster City, CA). Double-stranded PCR (ds-PCR) segments encompassing the entire control region (1,121 bp) were amplified with primers complementary to the light (np 15978–15997, 5′-CTACGCCAATCACTTTATTG-3′) and heavy (np 429–408, 5′-CTGTTAAAAGTGCATACC-GCC-3′) strands of the mtDNA, and then purified with Centricon-100 microconcentrators (Amicon, Inc., Danvers, MA) to remove excess primers and dNTPs. The purified ds-PCR fragments were sequenced with the Taq DyeDeoxy Termination Cycle Sequence Kit™ (Perkin Elmer, Oak Brook, IL) using primers complementary to the light (np 15978–16000, 5′-ACCATTAGCACCCAAAG-CTA-3′; np 16225–16244, 5′-CAACTATCA-CACATCAACTG-3′) and heavy (np 16421–16402, 5′-TGATTTCACGGAGGATGGT-3′) strands. The excess of DyeDeoxy™ terminators was removed from the completed sequencing reactions by using Centri-Sep™ columns (Princeton Separations, Adelphi, NJ), which were then run on 4% polyacrylamide/6 M urea/1× Tris-Borate-EDTA (TBE) gels. The resulting sequence data were collected and analyzed using the SEQED software included in the ABI 377A DNA Sequencer. Alignments and comparisons of the CR sequences were performed using the Sequencher 3.0 software tool (Gene Codes Corporation, Ann Arbor, MI).

Statistical analyses

Gene diversity. To quantify the amount of mtDNA diversity within each population irrespective of the phylogenetic relationships between different haplotypes, we estimated the gene diversity, or heterozygosity (h), of these groups from both RFLP haplotype and CR sequence data using equation 8.5 from Nei (1987),

$$h = \frac{(1 - \Sigma\, x_i^2) \cdot n}{n - 1}$$

where x_i is the frequency of each mtDNA haplotype and n is the number of haplotypes present in the population. The variance [V(h)] of each gene diversity estimate was also calculated using the formula

$$V(h) = \frac{2\propto [\Sigma\, x_i^3 - (\Sigma\, x_i^2)^2]}{n}$$

Probability of identity. Similarly, the relative genetic similarity of individuals within native Siberian populations was assessed by estimating the probability of identity (p) of any two mtDNAs, using the formula

$$p = \sum^{n} x_1 \cdot x_2$$

where n is the total number of haplotypes and x_1 and x_2 are the frequencies of any two unique haplotype within a population. In addition, the relative genetic similarity of Siberian populations was assessed by estimating the probability of identity (p) between them using the formula

$$p = \sum_{i,j}^{n} x_i \cdot x_j$$

where x_i and x_j are the frequencies of a shared mtDNA haplotype in populations i and j summed over the n haplotypes observed in the two populations (Nei, 1987).

Maximum likelihood estimates. The mean intra- and intergroup sequence divergence of the Koryaks and Itel'men as well as the other native Siberian (Torroni et al., 1993b; Starikovskaya et al., 1998) populations characterized by high resolution restriction analysis were estimated from RFLP haplotype data with the maximum likelihood (ML) procedure of Nei and Tajima (1983). Similar sequence divergence estimates were calculated for all of the major haplogroups present in eastern Siberians (A, C, D, G, Y, Z) and Native Americans (A–D) using the RFLP haplotype data from this and published studies (Torroni et al., 1992, 1993a,b, 1994a,b; Huoponen et al., 1997; Starikovskaya et al., 1998). When calculating the divergence times for intra- and intergroup variation as well as for indi-

vidual haplogroup variation, we used a mtDNA evolutionary rate of 2.2–2.9% per million years (MYR) (Torroni et al., 1994a).

Nucleotide diversity. The average nucleotide diversity within and between populations was estimated with the Sendbs program (N. Takezaki; http://iubio.bio.indiana.edu), which uses the method of Nei and Jin (1989) to estimate pairwise diversity values. In this analysis, diversity estimates were calculated from the CR sequences from Siberians (Torroni et al., 1993b; Starikovskaya et al., 1998; this study), east Asians (Horai et al., 1996), and Native Americans (Ward et al., 1991, 1993; Shields et al., 1993). Several different DNA distances were calculated from the diversity estimates, and standard errors (S.E.) of each value were obtained by bootstrapping over all sites using 500–1,000 replications, with the 95% confidence interval (C.I.) for the diversity and divergence estimates being calculated by using ± 2 S.E. In addition, neighbor-joining (NJ) (Saitou and Nei, 1987) trees were generated from the different genetic distances estimated from the nucleotide diversity values. Since all distances gave NJ trees which showed the same relationships among the populations being studied, the one generated from the Kimura two-parameter (1980) distance method is presented here.

Phylogenetic analyses

The evolutionary relationships between the complete haplotypes of the Koryaks, Itel'men, and other native Siberians (Torroni et al., 1993b) were inferred by parsimony analysis. Maximum parsimony (MP) trees were generated from the complete haplotype data through heuristic searches using the tree bisection and reconnection (TBR) branch-swapping algorithm with the random addition of taxa in PAUP (version 3.1.1) (Swofford, 1994). All MP trees were rooted from three African haplotypes, AF71 (Chen et al., 1995), TYPE-5, and HYPANC (Cann et al., 1987), which had the HpaI np 3592 site gain defining African macrohaplogroup L (Chen et al., 1995). After all runs, strict and 50% majority rule consensus trees were generated from the saved MP trees to determine the consistency of the branching arrangements.

Parsimony trees were also generated from the Siberian CR sequence data with the DNAPARS program in PHYLIP 3.572 (Felsenstein, 1994). To test the results of this analysis, all CR sequences were bootstrapped over 500 replicates using SEQBOOT and the resulting data files run in DNAPARS to generate unrooted parsimony trees. The parsimony trees were then used to generate strict and 50% majority rule consensus trees with CONSENSE, with the 50% majority rule consensus tree providing approximate bootstrap values for each branch of the tree (Felsenstein, 1994). Although DNAPARS produces a single tree from any particular data set, outgroup sequences can be used to root the tree, and the AF62 CR sequence (Chen et al., submitted) was used for this purpose.

Phylogenies of Siberian CR sequences were also inferred with the NJ method. This method was used because it is known to reconstruct correct phylogenetic trees with a high probability when analyzing closely related samples (Saitou and Imanishi, 1989). Rooted and unrooted NJ trees were generated from genetic distances estimated with models available in DNADIST in PHYLIP, including the Kimura two-parameter (Kimura, 1980), Jukes and Cantor (1969), and DNAML (Felsenstein, 1994) methods. The data sets used to generate these NJ trees included CR sequences from only Siberian populations (Torroni et al., 1993b; Starikovskaya et al., 1998; this study) and those from both Siberian and east Asian (Torroni et al., 1993b; Horai et al., 1996) populations, with the robustness of all trees being checked by bootstrapping using the algorithms (SEQBOOT and CONSENSE) available in PHYLIP (Felsenstein, 1994).

For population comparisons, NJ trees were constructed from ML values estimated with the Nei and Tajima (1983) method from high resolution RFLP haplotypes in Siberian populations (Fig.2). Similarly, NJ trees were constructed from genetic distances estimated from haplogroup frequencies in native Siberian and Asian populations analyzed by high resolution RFLP methods using the methods (Cavalli-Sforza and Edwards,

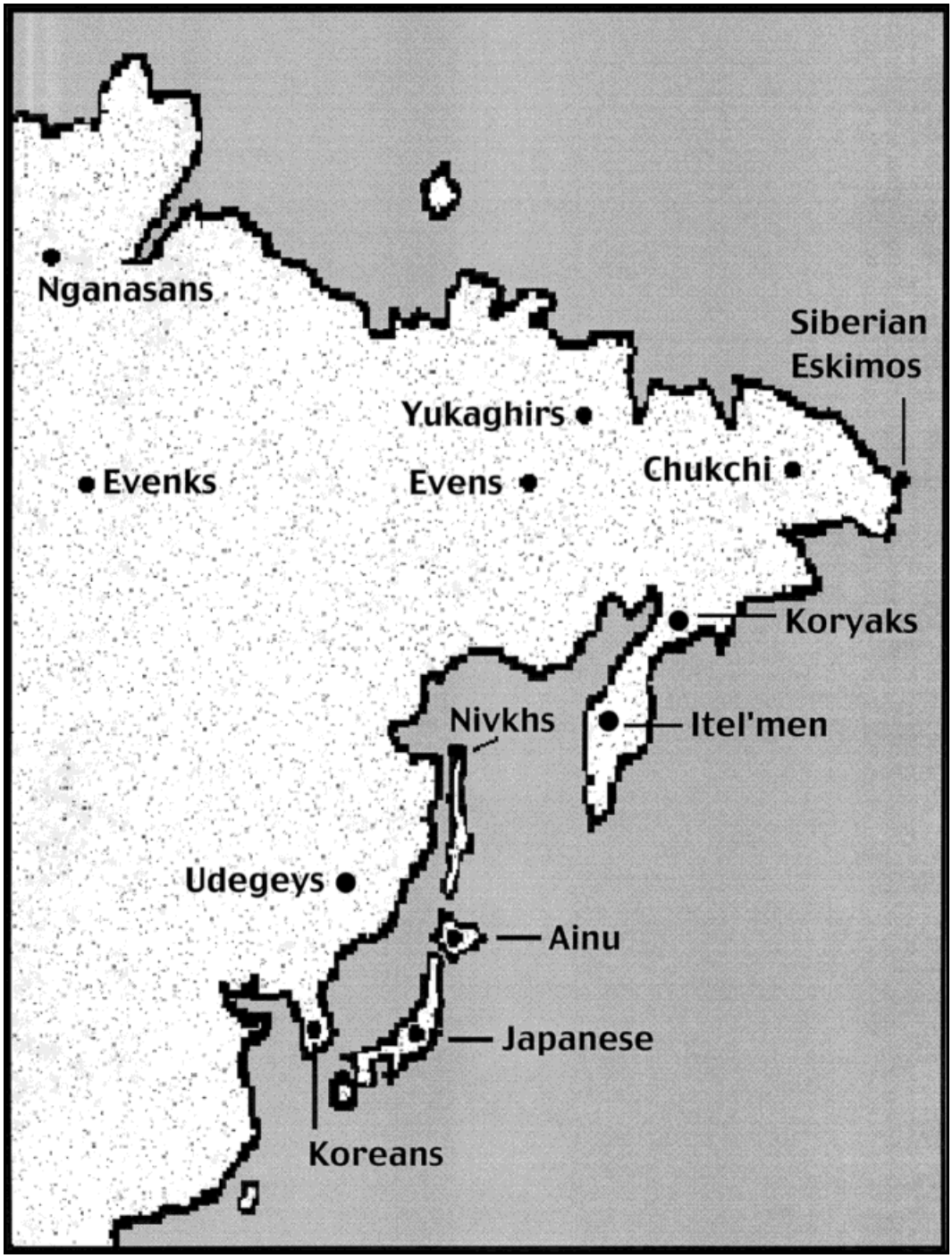

Fig. 2. A map of northeast Asia with the geographic locations of the Siberian and east Asian populations analyzed or compared in this study.

TABLE 1. RFLP haplotypes in Kamchatkan populations[1]

Haplogroup	Haplotype	Polymorphic restriction sites
A	SIB41	+663e
	SIB42	+663e, −5823a
C	SIB26	+10394c, +10397a, −13259o/+13262a, +16517e
	SIB29	−1715c, +10394c, +10397a, −13259o/+13262a, +15606a, +16517e
	SIB45	+10394c, +10397a, −13259o/+13262a
	SIB64	+3391e, +10394c, +10397a, −13259o/+13262a, +16517e
	SIB65	+8484a, +10394c, +10397a, −13259o/+13262a, +16517e
	SIB66	+1004h, +10394c, +10397a, −13259o/+13262a, +16517e
D	SIB40	−5176a, −10180l, +10394c, +10397a, +13717a, +14923c, +15437e
G	SIB08	+4830n/+4831f, +8198a, +10394c, +10397a, +16517e
	SIB35	+1588f, +4830n/+4831f, +8198a, +10394c, +10397a, +15494c, +16517e
	SIB36	+4830n/+4831f, +4923e, +8198a, +10394c, +10397a, +15494c, +16517e
	SIB37	+4830n/+4831f, +8198a, +10394c, +10397a, +15494c, +16517e
	SIB38	+4830n/+4831f, +8198a, +10394c, +10397a, +15494c
	SIB39	−3315e, +4830n/+4831f, +8198a, +10394c, +10397a, +15494c, +16517e
	SIB60	+4830n/+4831f, +8198a, +10394c, +10397a, +15494c, −16065g, +16517e
	SIB61	+1588f, +4830n/+4831f, +8198a, +10394c, +10397a, +15494c
Y	SIB01	+7933j, −8391e, +10394c, +16517e
Z	SIB43	+10394c, +10397a, +16517e
	SIB44	+10394c, +10397a, +11074c, +16517e
	SIB62	+6319i, +10394c, +10397a, +11074c, +16517e
	SIB63	−74171, +10394c, +10397a, +11074c, +16517e

[1] Haplotypes SIB01–SIB34 were previously defined in Siberians (Torroni et al., 1993b), and SIB46–SIB59 were defined in the Chukchi and Siberian Eskimos (Starikovskaya et al., 1998) but are not presented here. Polymorphic restriction sites are numbered from the first nucleotide of the recognition sequence according to the published sequence (Anderson et al., 1981). The restriction enzymes which detected variation are designated by the following single-letter code: a, *AluI;* c, *DdeI;* e, *HaeIII;* f, *HhaI;* g, *HinfI;* h, *HpaI;* i, *HpaII;* j, *MboI;* l, *TaqI;* n, *HaeII;* o, *HincII*. Sites separated by a diagonal line indicate either simultaneous site gains or site losses for two different enzymes or a site gain for one enzyme and a site loss for another because of a single common nucleotide substitution. These sites are considered to be only one restriction site polymorphism in the statistical analysis. All samples differ from the published sequence (Anderson et al., 1981) by the presence or absence of the following sites: −4769a, +7025a, +8858f, −13702e, −14199o, +14268g, and −14368g. In addition, the mtDNAs of four persons with maternal Even ancestry and paternal Koryak ancestry were analyzed and were found to have SIB08 (1), SIB37 (2), and SIB43 (1) haplotypes. Similarly, the mtDNAs of two persons with maternal and paternal Chukchi ancestry were analyzed and were found to have SIB26 (1) and SIB45 (2) haplotypes. None of these samples were included in the data for the Koryaks.

1967; Nei, 1972; Reynolds et al., 1983) available in GENDIST (Felsenstein, 1994). Each haplogroup was considered a distinctive "allele" since all of the RFLPs defining these haplogroups were essentially linked and distances calculated from the resulting "allele" frequencies. The alleles used to estimate genetic distances between populations included haplogroups A, B, C, D, F, G, Y, Z and "Other," with "Other" haplotypes being further classified into three subgroups, the first (I) being defined by -DdeI np 10394, -AluI np 10397, ±HaeIII np 16517, the second (II) by +DdeI np 10394, ± HaeIII np 16517, and the third (III) by +DdeI np 10394, +AluI np 10397, ± HaeIII np 16517.

RESULTS

mtDNA haplogroups in Kamchatkan populations

The high resolution RFLP analysis of 202 Koryak and Itel'men mtDNAs revealed a total of 22 distinct haplotypes defined by 48 polymorphic sites (Table 1), some of which had already been detected in Siberian populations. Three of the four haplogroups (A, C, and D) observed in Native American populations (Schurr et al., 1990; Torroni et al., 1992, 1993a) occurred in Kamchatkan groups, and these encompassed ~43% of all Koryak mtDNAs and 21% of Itel'men mtDNAs, with the majority of these belonging to haplogroup C. Consistent with previous studies of northeast Asian populations, Kamchatkan groups also lacked haplogroup B mtDNAs, suggesting these mtDNAs were never present in Paleoasiatic-speaking groups. In addition, none of the Koryak or Itel'men individuals exhibited mtDNAs from haplogroups typically seen in European populations (Torroni et al., 1994d, 1996), indicating that they had not experienced non-native gene flow through their maternal lineages.

Despite a third of their gene pool consisting of haplogroup A, C, and D mtDNAs, the Koryaks and Itel'men were not closely genetically related to Native American groups.

The only mtDNAs held in common between Kamchatkan and Native American populations were SIB41 (AM01) from haplogroup A and SIB26 (AM43) and SIB45 (AM32) from haplogroup C. All other haplotypes from these haplogroups were confined to either Chukotkan/Kamchatkan or Amerindian groups, indicating that they must have evolved in those respective geographic regions. In addition, the few haplogroup D mtDNAs in Koryaks (SIB40) were distantly related to those in Amerindian groups. Since the same haplotype appeared in the Chukchi and Siberian Eskimos (Starikovskaya et al., 1998) and related types have been detected at very low frequencies in the Koreans (Ballinger et al., 1992) and Japanese (Horai and Matsunaga, 1986), these mtDNAs apparently have a northeast Asian origin and distribution.

The majority of the Koryak (58%) and Itel'men haplotypes (80%) did not belong to haplogroups A, C, and D, and as such they could technically be defined as "Other" mtDNAs. However, we were able to further classify these putative "Other" mtDNAs into three additional clusters of related haplotypes through high resolution RFLP analysis (Table 1). In fact, this analysis clearly shows that most if not all Asian mtDNAs can be assigned to a haplogroup based on combined RFLP and CR sequence data, hence obviating the need for the category of the "Other" haplotype altogether.

The first cluster of "Other" haplotypes was defined by the combined HaeII np 4830 and HhaI np 4831 site gains and the linked DdeI np 10394 and AluI np 10397 site gains (hereafter called the DdeI/AluI sites). This cluster had previously been observed in Koreans and designated Asian haplogroup K (Ballinger et al., 1992) but was subsequently observed in Tibetans and renamed haplogroup G (Torroni et al., 1994c).

The second cluster of "Other" haplotypes was defined by the HaeIII np 8391 site loss and the MboI np 7933, DdeI np 10394, and HaeIII np 16517 site gains. This cluster had not previously been considered a haplogroup in eastern Siberians even though it encompassed haplotypes SIB01–SIB07 (Torroni et al., 1993b) and had also been detected in Koreans as AS105 from macrohaplogroup M in east Asians (Ballinger et al., 1992; Torroni et al., 1994c). It has now been renamed haplogroup Y, following the revised nomenclature begun in Torroni et al. (1993a,b) and continued in subsequent studies of mtDNA variation in different world populations (Torroni et al., 1994c,d, 1996; Chen et al., 1995).

The remaining Koryak and Itel'men haplotypes (SIB43, SIB44, SIB62, SIB63) belonged to a third cluster of mtDNAs. Of these haplotypes, SIB43 did not initially appear to belong to a well-defined haplogroup. In having only the DdeI/AluI sites and the HaeIII np 16517 site gain, SIB43 appeared to be identical with Asian haplotype AS118 (Ballinger et al., 1992) from Asian macrohaplogroup M (Ballinger et al., 1992; Torroni et al., 1994c). Haplotypes SIB44, SIB62, and SIB63 had the three RFLPs present in SIB43 but also showed an additional variant, the DdeI np 11074 site gain. This site gain is created by an A-to-G transition at np 11078 in the ND4 gene and converts an isoleucine (ATT) to valine (GTT). Consequently, these haplotypes constituted a new mtDNA lineage, designated haplogroup Z, with the relationship of SIB43 to these haplotypes being somewhat ambiguous based on RFLP data alone.

With respect to the DdeI np 11074 site gain itself, this polymorphism had not previously been detected in any haplotypes from other world populations except for AM83, which occurred in the South American Makiritare (Torroni et al., 1993a). AM83 had the four RFLPs present in SIB44 but differed from the latter by four additional polymorphisms, two of which (HinfI np 717 site gain and linked HaeII np 1622 and HhaI np 1623 site gains) were unique to AM83, and two (RsaI np 16049 site loss and HaeIII np 16517 site gain) being mutations which have arisen multiple times in different haplotypes from various world populations (Ballinger et al., 1992; Torroni et al., 1993a, 1994c, 1996; Chen et al., 1995; Starikovskaya et al., 1998). Because AM83 shared the RsaI np 16049 site loss with several haplogroup C mtDNAs in other Amazonian Indian tribes, it was suggested to be a haplogroup C mtDNA that had lost the characteristic markers of haplogroup C through a reversion mutation (Torroni et al.,

TABLE 2. mtDNA haplotype distribution in Kamchatkan populations[1]

Haplo-group	Haplo-type	Koryaks				Itel'men
		Aluitor	Karagin	Palan	Totals	
A	SIB41	2	0	2	4	0
	SIB42	1	1	2	4	3
C	SIB26	5	7	24	36	7
	SIB29	0	1	0	1	0
	SIB45	4	9	0	13	0
	SIB64	0	0	1	1	0
	SIB65	0	0	4	4	0
	SIB66	0	0	1	1	0
D	SIB40	2	0	0	2	0
G	SIB08	1	0	0	1	0
	SIB35	2	1	0	3	3
	SIB36	0	1	0	1	0
	SIB37	34	4	16	54	26
	SIB38	1	1	0	2	0
	SIB39	0	3	0	3	0
	SIB60	0	0	0	0	3
	SIB61	0	1	0	1	0
Y	SIB01	4	8	3	15	2
Z	SIB43	0	3	0	3	0
	SIB44	0	4	2	6	1
	SIB62	0	0	0	0	1
	SIB63	0	0	0	0	1
Totals		56	44	55	155	47

[1] The distribution of complete haplotypes among the three Koryak subgroups was significantly different, whether considering haplogroup (Fisher's exact test: $\chi^2 = 42.61$, $P = 0.0000$, d.f. = 10) or haplotype (Fisher's exact test: $\chi^2 = 90.59$, $P = 0.0000$, d.f. = 36) frequencies. In addition, the distribution of complete haplotypes between the Koryaks and Itel'men was statistically significant using both haplogroup (Fisher's exact test: $\chi^2 = 12.15$, $P = 0.0328$, d.f. = 5) and haplotype (Fisher's exact test: $\chi^2 = 46.62$, $P = 0.0035$, d.f. = 21) frequencies.

TABLE 3. Complete haplotype distribution in Maritime and Reindeer Koryaks[1]

Subgroup	N	Haplogroup					
		A	C	D	G	Y	Z
Maritime	89	5.6	31.5	1.1	43.8	7.9	10.1
Reindeer	54	5.6	40.7	1.9	37.0	14.8	0.0
"Koryak"	12	0.0	50.0	0.0	50.0	0.0	0.0
Totals	155	5.2	36.1	1.3	41.9	9.7	5.8

[1] Ethnicity was assigned to individuals based on self-identification as recorded from interview data. The category "Koryak" represents persons who were not certain of their ethnicity in terms of Maritime vs. Reindeer Koryak. This group was not compared against the Maritime and Reindeer Koryak subgroups in these chi-square tests. The assessment of subgroup differences did not show statistically significant differences in haplogroup distributions between the Maritime and Reindeers based on haplogroup frequencies (Fisher's exact test: $\chi^2 = 9.233$, $P = 0.1001$, d.f. = 5).

1993a). Subsequent phylogenetic analysis confirmed this interpretation, as the CR sequence for AM83 clustered with other haplogroup C mtDNAs from Native American populations (Schurr and Brown, unpublished). This result implied that SIB44 was not an ancestral form of AM83 which was brought across the Bering Strait by the ancient colonizers of the New World but instead that AM83 was a Native American haplogroup C mtDNA which had lost the characteristic markers of this mtDNA lineage in situ.

Haplotype distribution in Kamchatkan populations

An analysis of the distribution of RFLP haplotypes in the Koryaks and Itel'men provided a number of insights into the genetic relationships of the different Kamchatkan populations (Table 2). To begin with, the haplotype composition among the three territorial subgroups of Koryaks was markedly different. The Aluitor differed from the other two Koryak subgroups as well as from the Itel'men in having haplotype SIB40 from haplogroup D. In addition, the Aluitor subgroup lacked haplogroup Z haplotypes and had the highest frequency of haplogroup G mtDNAs. By contrast, SIB26 from haplogroup C was more common in the Palan subgroup, and novel haplotypes which derived from it (SIB64–66) were observed solely in this subgroup. A statistical assessment of this distribution[1] showed that it was significantly different, whether considering haplogroup or haplotype frequencies.

Interestingly, there were no significant differences in haplotypic composition between persons who were self-identified as either Maritime or Reindeer Koryak by ancestry irrespective of their village of origin (Table 3). Based on these results, Reindeer Koryaks appeared to be genetically synonymous with Maritime Koryaks, not a separate subgroup of this population, despite speaking a different dialect and practicing a different subsistence strategy. While these nonsignificant differences may reflect the fact that a certain proportion of the Koryak individuals sampled were not completely certain of their maternal ancestry in terms of Maritime or Reindeer Koryak ethnicity, it is more probable that they reveal the degree to which sedentary and nomadic groups

[1]All chi-square analyses were carried out with various methods available in StatXact 3 (CTEL Corporation, Cambridge, MA). Since most comparisons included three populations, chi-square analyses using Fisher's exact test were conducted using Monte Carlo estimates of *P* values, with a 99% confidence interval.

have become mixed in the past several centuries.

There were also noteworthy differences in the haplogroup distribution in the Koryaks of the Kamchatka peninsula relative to that of the Koryaks from northeastern Kamchatka, who had haplogroup A and D frequencies comparable to Reindeer Chukchi populations (Torroni et al., 1993b; Starikovskaya et al., 1998). The reason for this discrepancy lies in the source of samples for these populations. The previously analyzed "Koryak" population consisted of individuals from Middle Pakhachi and Achayvayam villages who were sampled as part of a study of conventional genetic markers in Chukchi populations (Sukernik et al., 1981, 1986). Extensive Chukchi admixture in these villages, if not a total replacement of the resident Koryaks resulting from prolonged Chukchi-Koryak wars in the nineteenth century (Bogoras, 1910), probably accounts for the difference in haplogroup composition of the Koryak subgroups. Consequently, the Middle Pakhachi-Achayvayam subgroup, originally classified as Reindeer Koryaks by Gurvich (1966), should instead be more properly considered Reindeer Chukchi, as suggested by Bogoras (1910).

When we consider both Kamchatkan groups, the mtDNA distribution in the Koryaks and Itel'men was also quite different (Table 2). While a number of haplotypes were shared between these two Paleoasiatic-speaking groups (SIB01, SIB26, SIB35, SIB37, SIB42, SIB44), the majority of these were the founding, or nodal, haplotypes for haplogroups C, G, Y, and Z (see Fig. 3). Otherwise, the Itel'men also lacked any unique haplotypes (SIB64–66) from haplogroup C and almost none from haplogroup G, whereas two unique haplogroup Z haplotypes (SIB62–63) occurred in this group. The extent of these genetic differences between the Koryaks and Itel'men was statistically significant whether considering haplogroup or haplotype frequencies.

The same extent of divergence was observed when the mtDNA variation in all three Paleoasiatic-speaking populations was assessed (Table 4). Chi-square analysis of the Chukchi, Koryaks, and Itel'men mtDNA distributions revealed statistically significant differences among them using either haplogroup (Fisher's exact test: $\chi^2 = 145.10$, $P = 0.0000$, d.f. = 10) or haplotype (Fisher's exact test: $\chi^2 = 228.2$, $P = 0.0000$, d.f. = 58) frequencies, while similar values were obtained for pairwise comparisons of these groups (data not shown). If we assume that their languages are closely related, these results could indicate that Paleoasiatic-speaking groups have undergone significant genetic differentiation since sharing a common origin in northeastern Siberia. Alternatively, these differences could suggest the separate origin and expansion of these populations in this region, with their linguistic affiliations reflecting the considerable language sharing which has taken place over the past several millennia. In the case of the Chukchi and the Koryaks, whose linguistic connection is more strongly supported, it appears that the Chukchi have become genetically distinctive from the Koryaks through considerable gene flow with Siberian Eskimos and perhaps other ethnic populations this region, such as the Yukagirs and Evens. Additional data from these latter groups will be necessary to clarify these interpretations.

Haplotype distribution in eastern Siberians

The recent studies of RFLP variation in Kamchatkan populations permitted a broader comparison of haplogroup distributions in eastern Siberian groups (Table 5). Aside from the Koryaks and Itel'men, haplogroup G mtDNAs were observed in the Chukchi but were absent in the Siberian Eskimos (Starikovskaya et al., in press). South of Kamchatka, haplogroup G mtDNAs were detected in 5.3% of the Nivkhs (Torroni et al., 1993b), 23.1% of the Koreans (Ballinger et al., 1992), and 7.5% of the Japanese (Horai et al., 1984; Horai and Matsunaga 1986). Unfortunately, it was not possible to accurately determine the frequency of haplogroup G mtDNAs in the Ainu from published data, as the previous RFLP analysis of Ainu mtDNAs (Harihara et al., 1988) did not use the enzymes which detect the characteristic RFLPs of haplogroup G

TABLE 4. MtDNA haplogroup distribution in Siberian and East Asian populations[1]

Language group / Population	n	Haplogroup frequencies (%)												Reference
											Other			
		A	B	C	D	E	F	G	Y	Z	I	II	III	
Siberian Yupik														
Siberian Eskimos	79	77.2	—	2.5	20.3	—	—	—	—	—	—	—	—	1
Paleoasiatic														
Chukchi	66	68.2	—	10.6	12.1	—	—	9.1	—	—	—	—	—	1
Koryaks	155	5.2	—	36.1	1.3	—	—	41.9	9.7	5.8	—	—	—	2
Itel'men	47	6.4	—	14.9	—	—	—	68.1	4.3	6.4	—	—	—	2
Isolated language														
Nivkhs	57	—	—	—	28.1	—	—	5.3	64.9	—	1.8	—	—	3
Tungusic														
Udegeys	45	—	—	17.8	—	—	—	—	8.9	—	28.9	—	44.4	3
Evenks	51	3.9	—	84.3	9.8	—	2.0	—	—	—	—	—	—	3
Koreans	13	7.7	7.7	—	23.1	7.7	15.4	15.4	7.7	—	—	7.7	7.7	4
Taiwanese Han	20	10.0	20.0	5.0	5.0	—	5.0	—	—	—	15.0	15.0	15.0	4
Tibetans	54	11.1	5.6	3.7	16.7	7.4	14.8	5.6	—	—	1.9	—	33.3	5

[1] "Other" haplotypes are those which do not belong to the haplogroups identified in this table but which may have different haplogroup affiliations. The mutational composition of these "Other" haplotypes was as follows: I, DdeI np 10394, −AluI np 10397, ±HaeIII np 16517; II, +DdeI np 10394, ±HaeIII np 16517; III, +DdeI np 10394, +AluI np 10397, ±HaeIII np 16517. The references cited in the table are as follows: 1, Starikovskaya et al. (1998); 2, this study; 3, Torroni et al. (1993b); 4, Ballinger et al. (1992); 5, Torroni et al. (1994c). For the three East Asian populations, there are some discrepancies between this table and the haplogroup frequencies published in Torroni et al. (1994c). In Torroni et al. (1994c), haplogroups A, B, C, D, and F corresponded to haplotype groupings H, D* + C, R, L, and A, respectively, from Ballinger et al. (1992), whereas haplogroups E and G were newly designated haplogroups. In being defined by the −HhaI np 7598, +DdeI np 10394, and +AluI np 10397 polymorphisms, haplogroup E may be equivalent to haplotype grouping G of Ballinger et al. (1992), while haplogroup G replaces haplotype grouping K from the same paper. In this table, haplotype grouping C from Ballinger et al. (1992) is removed from haplogroup B, with the haplotypes belonging to this mtDNA lineage being tallied in the Other II category, as a specific label has not yet been given to it. The reason for this separation is that the haplotypes from these two haplogroups are mutationally distinctive from each (Ballinger et al., 1992; Passarino et al., 1993) and clearly segregate into distinct clusters in MP trees of Asian mtDNAs (Fig. 2 [Ballinger et al., 1992]). Since these Other II haplotypes occur in both the Koreans and Taiwanese Han, they have been excluded from the haplogroup B column for these populations. In addition, one of the Korean mtDNAs (AS105 [Ballinger et al., 1992]) placed in the "Other" category by Torroni et al. (1994c) was reclassified as belonging to haplogroup Y since it was identical to SIB01; this change reduced the overall frequency of "Other" haplotypes in Koreans. Furthermore, the Korean haplotype (AS104 [Ballinger et al., 1992]) placed in haplogroup E by Torroni et al. (1994c) has both the HhaI np 7598 site loss from haplogroup E as well as the linked HaeII np 4830/+HhaI np 4831 site gains from haplogroup G. Hence, its phylogenetic status is ambiguous, although suggesting some sort of association between these two haplogroups. Similarly, one of the Taiwanese Han haplotypes (AS61 [Ballinger et al., 1992]) that was placed in haplogroup B by Torroni et al. (1994c) had the linked HincII np 12406/HpaI np 12406 site losses defining haplogroup F and the region V 9 bp deletion defining haplogroup B but lacked the HaeIII np 16517 site gain that is almost always present in haplotypes from the latter mtDNA lineage. Thus, its phylogenetic status is also uncertain. Consequently, AS104 and AS105 will require further sequence analysis to determine whether or not they have been placed in the correct haplogroup category. As for the remaining haplotype groupings of Ballinger et al. (1992), most of these (B, G, I, J, O, P, Q, S, T) appear to represent legitimate haplogroups which are present in East and Southeast Asian populations and which are now being reclassified due to these letter designations having been given to additional haplogroups that are present in other world populations (e.g., Torroni et al., 1994d, 1996). In the case of haplotype groupings E and F from Ballinger et al. (1992), these did not originally represent a single haplogroup but instead those mtDNAs whose general mutational characteristics included the +DdeI np 10394 and +AluI np 10397 polymorphisms. As such, they can be considered equivalent to Asian macrohaplogroup M, which is defined in the same way (Torroni et al., 1994c; Chen et al., 1995). Likewise, haplotype grouping D of Ballinger et al. (1992) represents a number of mtDNAs which probably belong to different haplogroups, including those from haplogroup B.

(HaeII np 4830 and HhaI np 4831 site gains).

By contrast, haplogroup Y mtDNAs were absent in most Siberian populations, including the Chukchi and Eskimos. However, these haplotypes represented the third most frequent haplogroup in the Koryaks (9.7%) and were present at polymorphic frequencies in the Itel'men (4.3%). In the lower Amur River region, this haplogroup is common in the Udegeys (8.9%) (Torroni et al., 1993b) and reaches its highest frequency in the Nivkhs of northern Sakhalin (64.9%). Although similar mtDNAs have also been observed at polymorphic frequencies in the Koreans (7.7%) (Ballinger et al., 1992), their presence in the Japanese is uncertain, as previous RFLP analyses of Japanese mtDNAs did not use enzymes which would detect all characteristic RFLP markers for this haplogroup (Horai et al., 1984; Horai and Matsunaga, 1986). SIB01 was the most common haplogroup Y haplotype in Siberian populations (Torroni et al., 1993b; this study), suggesting that it was the founding haplotype for this mtDNA lineage. Since derivative haplotypes have been seen only among the Nivkhs and Udegeys (Torroni et al.,

TABLE 5. Genetic diversity and probability of identity estimates for Siberian populations[1]

Population	N	n	Most common haplotype	Gene diversity (h ± S.E.)	Probability of identity within population (%)	Average probability of identity between population (%)	Ratio
Evenks	51	16	21.6	0.888 ± 0.001	47.2	3.62	13.04
Udegeys	45	10	28.9	0.843 ± 0.002	39.5	0.76	51.97
Nivkhs	57	11	45.6	0.732 ± 0.006	38.5	1.71	22.51
Koryaks	155	19	34.8	0.807 ± 0.000	41.8	7.75	4.86
Itel'men	47	9	55.3	0.739 ± 0.012	31.5	7.74	4.07
Chukchi	66	11	39.4	0.781 ± 0.004	40.6	4.46	9.10
Eskimos	79	12	36.7	0.805 ± 0.003	40.1	3.22	12.45

	Evenks	Udegeys	Nivkhs	Koryaks	Itel'men	Chukchi	Eskimos
Evenks	—	2	0	2	1	1	1
Udegeys	2.36	—	1	1	1	0	0
Nivkhs	0.00	2.01	—	2	1	0	0
Koryaks	5.11	0.43	4.44	—	6	3	3
Itel'men	3.22	0.19	1.96	23.50	—	1	1
Chukchi	2.29	0.00	0.00	3.52	1.58	—	6
Eskimos	0.28	0.00	0.00	1.00	0.19	19.22	—

[1] In the top panel, unbiased estimates of diversity and probability of identity estimates for eight native Siberian populations analyzed with the high resolution RFLP method are shown. N, the total number of individuals analyzed per population; n, the number of haplotypes observed in each population. The average probability of identity between populations is estimated as the average probability of identity between each population and the other seven populations. The ratio is the probability of identity within populations/probability of identity between populations. The data for the Evenks, Udegeys, and Nivkhs were taken from Torroni et al. (1993b), and those for the Chukchi and Siberian Eskimos were taken from Starikovskaya et al. (1998). In the bottom panel, the probability of identity estimates based on shared haplotypes between populations is shown. The numbers above the diagonal indicate the number of mtDNA haplotypes shared between populations, while those below the diagonal indicate the percent similarity of the populations. The population abbreviations are the same as those used in the Appendix.

1993b), haplogroup Y appears to have evolved in the lower Amur River region of southeastern Siberia.

As for haplogroup Z, these haplotypes seemed to have a geographic focus in Kamchatka. No comparable mtDNAs were seen in the Chukchi or Siberian Eskimos (Starikovskaya et al., in press), Siberian populations south of Kamchatka (Torroni et al., 1993b), or east Asian populations (Ballinger et al., 1992). The possible exceptions were the Tibetans (Torroni et al., 1994c) or Bornean aborigines (Ballinger et al., 1992), amongst whom AS118 appeared to be identical to SIB43, although additional CR sequencing will be necessary to confirm their lineal association with haplogroup Z mtDNAs. However, haplogroup Z mtDNAs were detected in the Evens (Table 1 footnote), suggesting that these mtDNAs may have originated in Tungusic-speaking populations.

Genetic diversity of Siberian populations

The extensive mtDNA diversity among eastern Siberian groups was further demonstrated with various statistical analyses of the RFLP haplotype data. The unbiased estimate of diversity (h) indicated that native Siberians were genetically heterogeneous while relatively similar in their overall level of diversity, with the Evenks being the most diverse (0.888) and the Nivkhs (0.732) being the least diverse (Table 5, top panel). These values are also generally reflected in both the number of distinct haplotypes and the frequency of the most common haplotype (MCH) in each population, with those having high frequencies of their MCH also showing the lower diversity values.

These findings were further confirmed with probability of identity (p) estimates. The p values for mtDNAs within populations were five to 40 times higher (31.5–47.2%) than those of mtDNAs between populations (0.76–7.75%), indicating a substantial degree of population differentiation of these aboriginal groups (Table 5, bottom panel). The level of differentiation was seen more starkly when the probability of identity between groups was estimated from only

TABLE 6. Genetic divergence estimates for native Siberian populations[1]

	Eskimos	Chukchi	Koryaks	Itel'men	Nivkhs	Udegeys	Evenks
Eskimos	*0.1105*	0.1133	0.1571	0.1626	0.1493	0.1559	0.1721
Chukchi	0.0063	*0.1128*	0.1456	0.1534	0.1449	0.1534	0.1678
Koryaks	0.0453	0.0326	*0.1132*	0.1139	0.1342	0.1417	0.1506
Itel'men	0.0530	0.0434	0.0030	*0.1087*	0.1444	0.1589	0.1635
Nivkhs	0.0419	0.0363	0.0254	0.0379	*0.1044*	0.1457	0.1576
Udegeys	0.0244	0.0208	0.0089	0.0283	0.0173	*0.1525*	0.1606
Evenks	0.0555	0.0500	0.0328	0.0478	0.0440	0.0230	*0.1228*

[1] Genetic distance estimates for native Siberian populations analyzed with the high resolution RFLP method using the maximum likelihood procedure (Nei and Tajima, 1983). The numbers (in italics) along the diagonal indicate intrapopulational divergences, the numbers above the diagonal indicate interpopulational divergences and the numbers below the diagonal indicates interpopulational divergences corrected for intrapopulational divergences.

shared haplotypes. These estimates showed very small levels of between-group similarity (0.00–5.11%), with at most only three mtDNA haplotypes being shared between any two populations, except the Chukchi and Siberian Eskimos and the Koryaks and Itel'men, who both shared six different haplotypes.

Interestingly, both the Koryak populations and the Itel'men shared one to three haplotypes with the five other Siberian ethnic groups (Chukchi, Eskimos, Udegeys, Nivkhs, Evenks) analyzed. These apparent genetic affinities with other Siberian groups are also shown in the ratio of within-group to between-group identity values, which are lowest for the Koryaks and Itel'men. However, the majority of the mtDNAs shared between these populations were the putative founding haplotypes for haplogroups A, C, G, Y, and Z rather than more recently derived ones, which are usually population- or region-specific. Hence, while showing genetic affinities with both the Evenks and Amur River populations, the Koryaks and Itel'men may be somewhat more genetically distant from the other groups than implied by these estimates.

The maximum likelihood (ML) method of Nei and Tajima (1983) showed similar levels of diversity for eastern Siberian populations. Both Koryaks (0.113%) and Itel'men (0.108%) had intrapopulational divergence values which were comparable to those obtained for the other native Siberian populations analyzed by similar methods (Table 6). The lowest value for these groups was seen in the Nivkhs (0.104%), who had a predominance of two haplotypes (SIB01, SIB10), while the highest occurred in the Udegeys (0.153%), who exhibited a set of unique haplotypes (SIB21–SIB25) relative to those of the other groups (Torroni et al., 1993b). Furthermore, the close genetic affinities of the Koryaks and Itel'men were shown by their extremely small corrected interpopulational value, one approached only by the Chukchi-Siberian Eskimo value.

The ML values for Siberian population were also found to be higher than comparable estimates for most Native American populations (Torroni et al., 1994a). This difference could imply a greater antiquity of eastern Siberian populations relative to their New World counterparts. However, they may also reflect the recent hybridization of both interior and Pacific Siberian populations and the concomitant acquisition of mtDNA lineages from more southerly ethnic groups. These genetic influences can be seen by the presence in the Evenks of SIB20, a mtDNA haplotype belonging to haplogroup F (Torroni et al., 1993b), which commonly occurs in Tibetan (Torroni et al., 1994c) and southeast Asian (Ballinger et al., 1992) populations.

Phylogenetic analysis of Siberian haplotypes

The results of the MP analysis of Koryak, Itel'men, and other native Siberian RFLP haplotypes recapitulate the trends seen in the statistical analyses of these data (Fig. 3). With few exceptions, the shared haplotypes amongst these groups were the nodal, or putative founding, haplotypes for each haplogroup. The remaining shared haplotypes were usually common to geographically proximate populations, such as the Koryaks and Itel'men (e.g., SIB35 from haplogroup

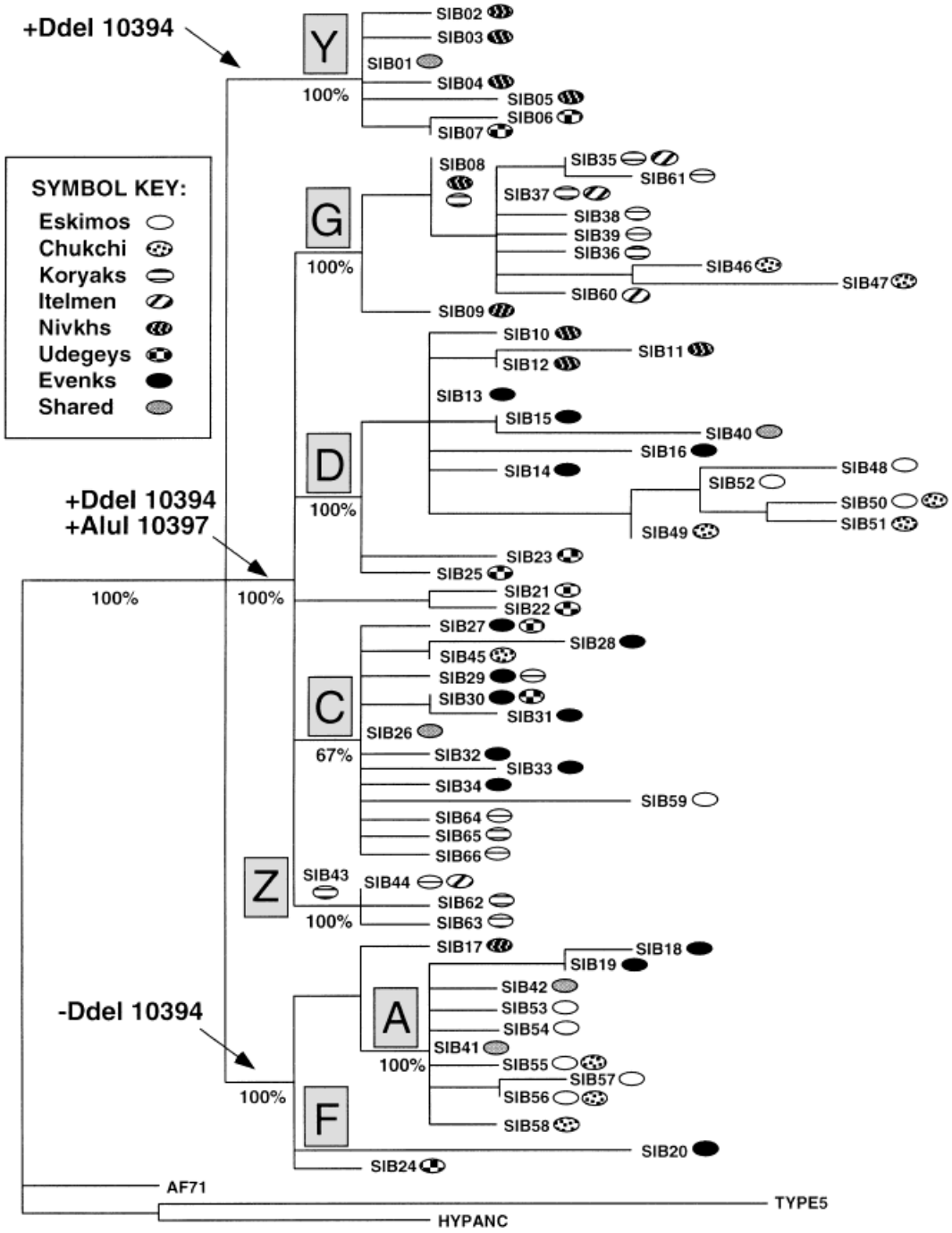

Fig. 3. An MP tree of Koryak, Itel'men, and other eastern Siberian RFLP haplotypes. The TBR tree is 119 steps in length, has a Consensus index (C.I.) of 0.765 and an retention index (R.I.) of 0.852, and represents one of 120 MP trees that were generated by the heuristic search irrespective of the number of MAXTREES specified. The mtDNA haplogroups observed in native Siberian populations are indicated by the black capital letters in shaded boxes, while haplotypes appearing in each population are identified by ellipses specified in the symbol key. For haplotypes shared between two populations, the appropriate ellipses for each population have been positioned behind the haplotype number; those appearing in more than two populations are indicated as shared (see Table 2). the African haplotypes used as outgroups to root this tree included AF71 (Chen et al., 1995) and HYPANC and Type-5 (Cann et al., 1987). The numbers located under the major branches of the MP tree represent the percent support for each branch observed in the 50% majority rule consensus tree.

G), or those with known linguistic affiliations, such as the Evenks and Udegeys (e.g., SIB27). In addition, all haplogroup branches were very highly supported in the 50% majority rule consensus tree, the weakest being haplogroup C. The same pattern was observed when Siberian haplotypes were analyzed with those from other Asian populations (Koreans, Taiwanese Han, Tibetans) (Ballinger et al., 1992; Torroni et al., 1994c) (results not shown), with the most divergent haplotypes from each haplogroup being largely population-specific and located at the terminal positions of the branches, and population-specific clusters occurring within some of these haplogroups (e.g., Koryaks and haplogroup G, Tibetans and haplogroups D and G).

When the ML estimates were used to construct NJ trees, the Paleoasiatic-speaking groups were split into two separate branches, one representing Chukotka and the other Kamchatka (Fig. 4). In addition, the Nivkhs branched off close to the Kamchatkan groups, while the Udegeys and Evenks formed a separate branch between the Nivkhs and Chukotkan groups. Nearly identical populational associations were seen in the NJ tree based on genetic distances estimated from haplogroup frequencies in Siberian and Asian populations (Fig. 5), with Koreans and Nivkhs clustering together between the Kamchatkan and Chukotkan populations. These associations among Siberian groups were concordant with the trends seen in the gene diversity and probability by identity estimates for the same groups (Tables 5, 6).

The centrality of the Koreans, Tibetans, and Taiwanese Han in Figure 5 results from each of these populations having at least four of the main haplogroups (A, C, D, F, G, Y, Z) which collectively appear among the populations shown in this tree. As such, this distribution reveals the genetic influences of both northern and southern Asian groups on these populations. On the other hand, the location of the Udegeys in this tree is probably due to their having a high frequency of "Other" haplotypes bearing the DdeI/AluI sites, which the Tibetans and Taiwanese Han also possess (Table 4), although the Udegey haplotypes are distinctive from those present in the other two populations. Likewise, the position of the Evenks likely reflects their sharing haplogroup C mtDNAs with both Siberian and Asian groups and having haplogroup F mtDNAs in common with central-east Asian groups.

CR sequence variation in Kamchatkan populations

The sequencing of the CR of Koryak and Itel'men mtDNAs provided a much more detailed picture of genetic variation in Paleoasiatic-speaking populations. Overall, this analysis revealed a total of 53 different CR sequences, as defined by 54 variable nucleotide positions, in Kamchatkan populations, most of which had not previously been observed (Table 7). The greatest sequence diversity was observed among haplogroup C and G mtDNAs, while a number of unique CR sequences were associated with haplogroups A, D, Y, and Z. In addition, a considerable degree of substructure within individual haplogroups was detected, allowing a finer discrimination of lineal associations between populations occupying northeast Siberia and the New World.

Haplogroup A. The CR sequences from Koryak haplogroup A mtDNAs which belonged to SIB41 had the np 16223T, np 16290T, np 16319A, and np 16362C mutations which define the sequence motif for this haplogroup in Asia and the Americas. In addition, these CR sequences had the np 16111T mutation. This polymorphism has been observed in nearly all haplogroup A mtDNAs of the Chukchi, Eskimos, Na-Dené Indians, and Amerindians (Torroni et al., 1993a; Forster et al., 1996; Starikovskaya et al., 1998) but is absent in those from Asian populations (Torroni et al., 1993b; Kolman et al., 1996). This pattern indicated that the np 16111T mutation arose in the earliest inhabitants of Beringia, who later gave rise to the ancestral Native American population(s) in which haplogroup A evolved and that the Koryaks have retained some haplogroup A mtDNAs which are related to other Chukotkan and New World populations.

The np 16192T transition was also observed in the Koryak haplogroup A CR se-

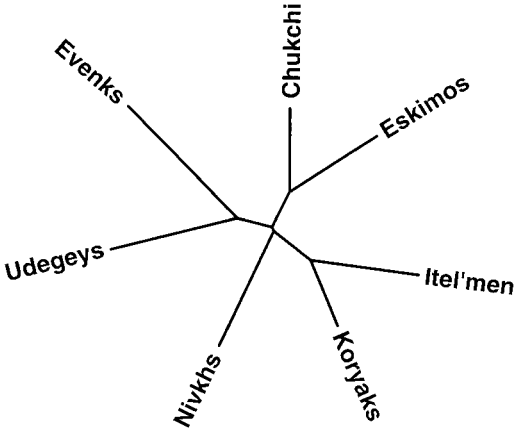

Fig. 4. An NJ tree based on ML estimates for Siberian populations analyzed by high resolution analysis (see Table 6).

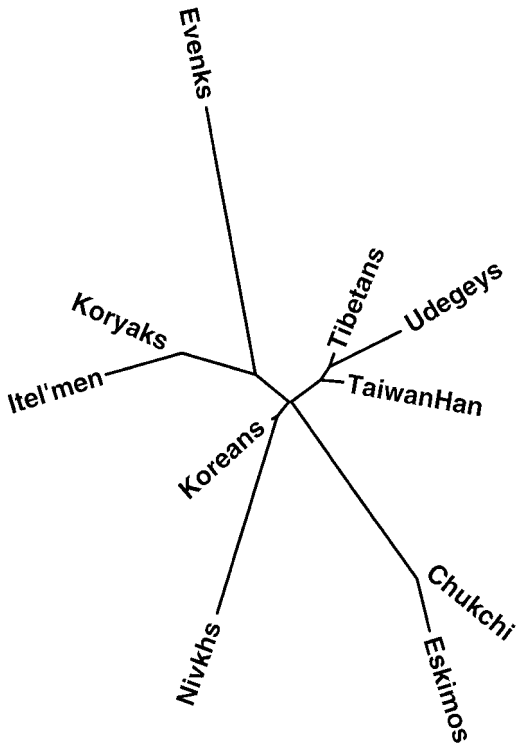

Fig. 5. An NJ tree based on genetic distances estimated with the Reynolds et al. (1983) method in PHYLIP (Felsenstein, 1994); the frequencies of the haplogroup "alleles" used to estimate the genetic distances for these populations are shown in Table 4.

quences which exhibited the 16111T mutation. This mutation had previously been observed in similar mtDNAs from the Chukchi, Siberian and North American Eskimos, Canadian Dogrib, and Navajo (Shields et al., 1993; Torroni et al., 1993b; Starikovskaya et al., 1998) and represents a major sublineage that is largely confined to the north Pacific Rim. The distribution of the 16192T mutation implies that it arose among the most recent common ancestors of the Chukchi, Eskimo-Aleuts, and Na-Dené Indians exclusive of the progenitors of Amerindian populations. Thus, the presence in the Koryaks of mtDNAs with this mutation likely reflects the preservation of ancient Beringian haplotypes in Kamchatkan groups.

In addition, both the Itel'men and Koryaks had haplogroup A mtDNAs which lacked both the 16111T and the 16192T polymorphisms as well as the 16362C mutation typically seen in haplogroup A mtDNAs in the New World. All of these were associated with haplotype SIB42, the only haplogroup A haplotypes appearing in the Itel'men. While it is possible that one or both of the Beringian mutations were lost from these CR sequences, they probably represent mtDNAs from haplogroup A which never experienced these nucleotide substitutions, as is the case for all similar Asian haplotypes (Torroni et al., 1993b; Kolman et al., 1996). The latter explanation is supported by the presence in these mtDNAs of the np 16242T mutation, which is not seen in comparable haplotypes from other circumpolar or New World populations (Ward et al., 1991, 1993; Shields et al., 1993; Torroni et al., 1993a).

Haplogroup C. Two distinct clusters of haplogroup C CR sequences were observed in both Kamchatkan populations (Table 8). The first encompassed CR sequences #05–12 and exhibited all of the polymorphisms which define haplogroup C mtDNAs in both Siberian and Native American populations, including the np 16223T, np 16298C, and np 16327T transitions (Torroni et al., 1993a,b). In addition, the majority of the CR sequences belonging to this cluster (#07–09) also exhibited the np 16124C transition, and many also had the np 16318T transversion, both of which do not appear in haplogroup C

TABLE 7. HVS-I sequences in Koryaks and Itel'men

Haplo-group	Haplo-type	HVS-I sequence number	Polymorphic nucleotides 00000111111111111222222222222222222222333333333333335555 11369112227788990122344666678899999990111112245666112 2 78893194692359277834123015637801234684016897822267395 7
		CAM	TTACTCATTGTCCTCCACCTTCTCCACGCTCCCACCTTGTAAGCCTTTCATTAC
A	SIB41	01	-----T--------T---T-----------T-----------A----C------
	SIB41	02	-----T--------T---T-----------T--------C--A---CC------
	SIB42	03	------------------T--T--------T-----------A----------
	SIB42	04	----C-------------T--T--------T-----------A----------
C	SIB29	05	---------A--------T-----------------C------T-------C--
	SIB26	06	------------------T-----------------C------T-------C--
	SIB26	07	-------C----------T-----------------C----T-T-------C--
	SIB64	07	-------C----------T-----------------C----T-T-------C--
	SIB45	08	-------C----------T-----------------C----T-T----------
	SIB26	09	-------C----------T-----------------C----T-T------CC--
	SIB26	10	-------C--------------------------C------T-------C--
	SIB26	11	-------C--------------------------C------T-----G-C--
	SIB26	12	-------C----------T-----------------C------T-------C--
	SIB26	13	----C--------C----T-----T----C------C--------------C--
	SIB65	13	----C--------C----T-----T----C------C--------------C--
	SIB66	13	----C--------C----T-----T----C------C--------------C--
	SIB26	14	-C--C--------C----T-----T----C------C--------------C--
	SIB26	15	----C--------C----T-----T-T--C------C--------------C--
	SIB26	16	----C--------C--------------------C--------------C--
	SIB26	17	----C--------C----T----TT----C------C--------------C--
	SIB26	18	----C--------C-T--T-----T----C------C--------------C--
	SIB26	19	-------------C----T-----T----C----------C-----------C--
	SIB26	20	------------------T-----T----C------C--------------C--
	SIB26	21	----C--------C----T----------C------C--------------C--
	SIB26	22	----C--------C----T-----T----C--------------------C--
	SIB26	23	-------------C----T-----T-T--C--------------------C--
D	SIB40	24	----C------T------T----------------------A----C------
G	SIB08	25	C--------A--------T--------------------------------C--
	SIB37	25	C--------A--------T--------------------------------C--
	SIB37	26	C---C----A--------T--------------------------------C--
	SIB39	26	C---C----A--------T--------------------------------C--
	SIB37	27	C---C----AC-------T------G-------------------------C--
	SIB37	28	C--------AC-------T------G-------------------------C--
	SIB37	29	C---C----A--------T------------T-------------------C--
	SIB37	30	C---C----A--------T--------------------------------C--
	SIB35	31	C---C----A--------T--------------------C-----------C--
	SIB61	32	C---C----A--------T--------------------C--------------
	SIB35	33	C--------A--------T--------------------C-----------C--
	SIB38	34	C--------A--------T--------------------C--------------
	SIB60	35	C--TC----A--------T--------------------------------C--
	SIB37	36	C---C-------------T--------------------------C-----C--
	SIB37	37	C--------A--------T--------------C-----------------C--
	SIB37	38	C---C----A--------T--------------C-----------------C--
	SIB37	39	C--------A--------T--------------------------------C-T
	SIB37	40	C-----G--A--------T--------------------------------C--
	SIB37	41	C--------A--------T--------A-----------------------C--
	SIB37	42	C---C----A--------T--------------------------------CG-
	SIB37	43	C---C----A------G-T--------------------------------C--
	SIB36	44	C---C-----------G-T--------------------------------C--
	SIB37	44	C---C-----------G-T--------------------------------C--
	SIB37	45	C---C-----------G-T-------------T-----------------C--
	SIB37	46	C--------A--------T--------------------------C---C--
Y	SIB01	47	--------C----C------C-----T-----------------------C--
	SIB01	48	--------C----C----T-C-----T-----------------------C--
	SIB01	49	--------C----C------C-----T-T-----------G---T------C--
	SIB01	50	--------C----C------C-----T-T---------C-G---T------C--
	SIB01	51	--------C----C------C---T-T-----------------T------C--
Z	SIB43	52	---------A--T-----TC---T------------C--------------C--
	SIB44	52	---------A--T-----TC---T------------C--------------C--
	SIB62	53	---------A--TC----TC---T------------C--------------C--
	SIB63	53	---------A--TC----TC---T------------C--------------C--

TABLE 8. Distribution of HVS-I sequences in Kamchatkan populations[1]

Haplogroup	CR sequence	Koryaks				Itel'men
		Aluitor	Karagin	Palan	Total	
A	01	0	0	2	2	0
	02	2	0	0	2	0
	03	1	1	2	4	0
	04	0	0	0	0	3
C	05	0	1	0	1	0
	06	0	0	6	6	0
	07	4	1	4	9	0
	08	4	8	0	12	0
	09	0	1	0	1	0
	10	0	0	2	2	0
	11	1	0	0	1	0
	12	0	0	1	1	0
	13	0	0	12	12	3
	14	0	1	0	1	0
	15	0	0	0	0	2
	16	0	0	1	1	0
	17	0	1	0	1	0
	18	0	0	1	1	0
	19	0	0	0	0	1
	20	0	1	0	1	0
	21	0	1	0	1	0
C	22	0	0	1	1	0
	23	0	0	1	1	0
D	24	2	0	0	2	0
G	25	21	0	2	23	2
	26	4	2	2	8	3
	27	0	0	2	2	0
	28	0	0	4	4	0
	29	0	0	0	0	4
	30	0	0	0	0	3
	31	2	1	0	3	1
	32	0	1	0	1	0
	33	0	0	0	0	2
	34	1	0	0	1	0
	35	0	0	0	0	3
	36	1	0	0	1	0
	37	0	1	0	1	0
	38	0	1	0	1	0
	39	1	2	0	3	0
	40	2	0	0	2	0
	41	1	0	0	1	0
	42	0	0	2	2	0
	43	0	0	0	0	3
G	44	4	1	4	9	10
	45	0	0	0	0	1
	46	1	0	0	1	0
Y	47	4	3	3	10	1
	48	0	0	0	0	1
	49	0	1	0	1	0
	50	0	1	0	1	0
	51	0	1	0	1	0
	52	0	6	2	8	1
	53	0	0	0	0	2
Total		56	37	54	147	46

[1] Certain samples did not sequence well and thus were not included in the totals per population and/or ethnic subgroup; these included seven from the Karagin Koryaks and one from the Palan Koryaks as well as one from the Itel'men. These omissions are the source of the discrepancy in the sample sizes between this table and Table 2.

mtDNAs from other Siberian or Native American tribes (Ward et al., 1991, 1993; Shields et al., 1993; Torroni et al., 1993a,b). Similar CR sequences were seen in some haplogroup C mtDNAs from the Chukchi and Siberian Eskimos (Starikovskaya et al., 1998), implying that these mtDNAs are common to all Paleoasiatic-speaking groups and perhaps also Eskimoan populations, in which haplogroup C occurs at low frequen-

cies (Merriwether et al. 1994; Starikovskaya et al., in press).

The first cluster was further associated with several different RFLP haplotypes, including SIB26, SIB29, SIB45, and SIB64. Because these SIB45 haplotypes differed from equivalent haplotypes in Native American populations (AM32) (Torroni et al., 1992, 1993a,b) by having the 16124C and 16318T mutations, they appear to have arisen in Kamchatkan groups independent of putatively identical haplotypes in New World groups, with both sets of SIB45 mtDNAs becoming differentiated from SIB26 through the loss of the HaeIII np 16517 site, a known hypermutable restriction site (Ballinger et al., 1992; Torroni et al., 1994c, 1996; Chen et al., 1995). Consequently, Siberian SIB45 mtDNAs cannot be considered founding haplogroup C haplotypes for Native American populations.

The second cluster encompassed CR sequences #13–23 but differed from the first by having the np 16093C, np 16189C, np 16261T, and np 16288C transitions and by lacking the 16318T and 16327T mutations. As seen in other Asian and Native American mtDNAs, the 16189C transition creates a homopolymeric stretch of Cs within a 14 bp hypervariable domain (np 16180–16193), which typically results in the insertion of an additional one or more Cs (Horai and Hayasaka, 1990; Horai et al., 1993). CR sequences from this cluster also occurred in the Chukchi but not in the Siberian Eskimos (Starikovskaya et al., 1998) or distantly related Asian and Native American groups on both sides of the Bering Strait (Shields et al., 1993; Ward et al., 1993; Torroni et al., 1993a,b), and thus appeared to be unique for Paleoasiatic-speaking populations. This cluster was also associated with the founding haplotype, SIB26, as well as the unique haplotypes SIB65 and SIB66. The presence of at least two distinct CR clusters within haplogroup C and the association of both of them with the founder haplotype, SIB26, implied a considerable degree of divergence of this mtDNA lineage in Asia. Furthermore, since CR sequences from the first and the second clusters differed on average by five or six mutations, they could represent multiple expansions of haplotypes from the same mtDNA lineage in northeastern Siberian groups at different times.

Haplogroup D. Only one type of CR sequence from haplogroup D was observed in the Koryaks, and it was linked with haplotype SIB40. This CR sequence (#24) was defined by five different sequence polymorphisms, most of which appeared in one or more of the other haplogroups present in the Kanchatkan populations (16093C, 16223T, 16319A, and 16362C transitions), and had one distinguishing polymorphism, the np 16173T transition. The same CR sequence also appeared in four Siberian Eskimos and Chukchi (Starikovskaya et al., 1998) and had previously been observed in Alaskan Inupik Eskimos (CR lineage #58) (Shields et al., 1993). However, it has not been found in other Siberian and Native American populations (Shields et al., 1993; Torroni et al., 1993a).

Haplogroup G. Haplogroup G showed the greatest diversity of CR sequences relative to the other haplogroups. In general, the CR sequences for haplogroup G were defined by a set of four different sequence polymorphisms, including the np 16017C, 16093C, np 16129A, and 16223T transitions. Along with the linked HaeII np 4830 and HhaI np 4831 site gains, the 16017C transition clearly distinguished haplotypes from this haplogroup from those of other Siberian or Asian mtDNA lineages. Similar CR sequences were also seen in the Chukchi, who exhibited a limited number of haplogroup G mtDNAs (Starikovskaya et al., 1998). These data indicated that haplogroup G mtDNAs were part of the genetic makeup of ancient Paleoasiatic-speaking populations and that this mtDNA lineage has undergone a considerable degree of genetic diversification since being brought to northeastern Siberia.

Haplogroup Y. The CR sequences for this haplogroup exhibited a set of nucleotide polymorphisms which were unique to this mtDNA lineage, including the np 16126C, 16189C, np 16231C, np 16266T, and 16519C transitions. One cluster within this haplogroup was defined by only these mutations, while another also had the np 16287T, 16316G, and 16328T mutations which do

not appear in the other Siberian haplogroups. Since both of these sublineages appeared in SIB01 haplotypes from the Koryaks and Itel'men, the founding haplotype for this haplogroup, they were probably part of the ancestral pool of Kamchatkan groups.

Haplogroup Z. The CR sequences for haplogroup Z also exhibited a set of nucleotide polymorphisms unique to this mtDNA lineage, including the 16129A, 16185T, 16223T, 16224C, 16260T, 16298C, and 16519C mutations. In addition, SIB62 and SIB63 had the 16189C polymorphism which in this case did not create a homopolymeric stretch of Cs. More importantly, CR sequences with these mutations were seen in both haplotypes SIB43 and SIB44. This finding confirmed that SIB43 belonged to haplogroup Z and suggested that it had lost the DdeI np 11074 site gain seen in the other haplotypes from this mtDNA lineage. Because haplogroup Z mtDNAs appeared in both the Koryaks and Itel'men, they too were likely part of the ancestral pool of Kamchatkan groups.

Distribution of CR sequences in Koryaks and Itel'men

As can be seen in Table 8, a number of CR sequences from haplogroups A, C, G, Y, and Z were shared amongst all Koryak subgroups (#03, #07, #26, #44, #47). Their ubiquity amongst the Koryaks and high frequency relative to other CR sequences suggested that they represented the founding mtDNAs for this Paleoasiatic-speaking group. The high frequency of CR sequence #25 in the Aluitor Koryaks (37.5%) suggested its origin in this subgroup and its spread to the Palan Koryaks and Itel'men through gene flow, with the Karagin Koryaks possibly missing this CR sequence due to drift effects. In addition, the Aluitor and Karagin Koryaks were found to share CR sequences #08, #31, and #39, a result which supported their known close linguistic association. On the other hand, the Karagin and Palan Koryaks shared CR sequence #52, suggesting some differences between them and the Aluitor Koryaks. Aside from these common or shared types, the majority of CR sequences occurred in one Koryak subgroup or another, with many showing village specificity (e.g., within the Palan subgroup, #06 and #13 were detected only in persons born in Voyampolka).

A more restricted number of CR sequences from haplogroups G, Y, and Z were shared between the Koryaks and Itel'men. Both Paleoasiatic groups had CR sequences #25, #26, #31, and #44 from haplogroup G, as well as #47 from haplogroup Y and #52 from haplogroup Z. Because of their prevalence in each population, these sequences are likely the founding types for their respective haplogroups. In addition, the np 16207G mutation defined a set of haplogroup G mtDNAs (#43–45) present in both the Koryaks and Itel'men. Its high frequency in the Itel'men (30.4%) relative to the Koryaks (6.1%) suggested that this set may have originated in this population and spread to the Koryaks through gene flow. The only other shared CR sequence (#13) occurred in the Palan Koryaks and Itel'men, a distribution which might reflect the southward expansion of Reindeer Koryak groups into traditional Itel'men territory (Jochelson, 1908). Otherwise, all remaining CR sequences were population-specific for either of the two Kamchatkan groups.

Another intriguing finding was the nonuniform distribution of the CR clusters from haplogroup C in the Koryaks and Itel'men. The cluster defined by the 16124C-16223T-16298C-16318T-16327T motif (#05–12) was distributed across all three Koryak subgroups at more or less the same frequency but was completely absent from the Itel'men. In contrast, the cluster defined by the 16093C-16189C-16261T-16288C motif (#13–23) was present in only the Karagin and Palan Koryaks as well as amongst the Itel'men. Such a pattern suggested that mtDNAs from the first cluster were brought to the Kamchatka peninsula by ancestral Koryak groups, whereas those from the second may already have been present in the Itel'men and subsequently acquired by the southernmost Koryak groups through intermarriage with Itel'men populations.

When CR sequence diversity in all Paleoasiatic-speaking groups was examined, striking differences between Chukotkan and Kamchatkan populations were observed

TABLE 9. CR sequence diversity in northeast Siberian populations[1]

Population	N	n	Number of unique types (% individuals)	Gene diversity ($h \pm$ S.E.)	Probability of identity within populations (%)	HVS-I sequences shared between populations					
						A #01	C #07	C #13	D #24	G #25	G #26
Eskimos	77	12	3 (3.9)	0.819 ± 0.008	22.6	37	1	—	4	—	—
Chukchi	65	19	23 (35.4)	0.883 ± 0.001	42.8	19	1	4	1	2	2
Koryaks	147	41	70 (47.6)	0.945 ± 0.000	36.6	2	9	12	2	23	8
Itel'men	46	19	26 (56.5)	0.931 ± 0.000	43.1	—	—	3	—	2	3

[1] N, number of mtDNAs subjected to CR sequencing; n, number of distinct sequences observed in each population. The Kamchatkan HVS-1 sequences enumerated in the table are equivalent to those present in Chukotkan populations in the following way: KAM01 = CHU01, KAM07 = CHU17, KAM13 = CHU19, KAM24 = CHU13, KAM25 = CHU20, and KAM26 = CHU21 (KAM, Kamchatka; CHU, Chukotka) (Starikovskaya et al., 1998; this study). The number of mtDNAs with each CR sequence is specified by population. The letter above each CR sequence number indicates the haplogroup to which each one belongs.

(Table 9). As a whole, the Koryaks and Itel'men showed a greater frequency of unique types than did the Chukchi and Siberian Eskimos. The latter population had a dramatically lower frequency of these unique types due to sharing ten different CR sequences with the Chukchi, including ones identical to #01, #07 and #24 from this study (Starikovskaya et al., 1998). Interestingly, nearly all of these CR sequences were the same set shared between the two Kamchatkan populations, with only #01 from haplogroup A and #24 from haplogroup D being more highly frequent in Chukotkan groups. Although their strong genetic similarity to Siberian Eskimos suggested that the Chukchi might have acquired haplogroup C and G mtDNAs from Koryak subgroups through recent gene flow, the fact that all Chukchi subdivisions analyzed for mtDNA variation possessed haplotypes from both haplogroups (Torroni et al., 1993b; Starikovskaya et al., 1998; this study; Schurr et al., unpublished data) argues that these mtDNAs were part of the ancestral gene pool for all Paleoasiatic-speaking groups.

Phylogenetic analysis of Siberian CR sequences

When these CR sequences were subjected to phylogenetic analysis, all six of the major haplogroups present in northeastern Siberians formed distinct branches in the resulting NJ tree (Fig. 6) irrespective of whether or not African mtDNAs were used as outgroups. The same overall pattern was obtained through parsimony analysis with DNAPARS (results not shown). In addition, the substructure within haplogroups A, C, and D noted in the CR sequence data was clearly revealed in this phylogeny. Within haplogroup A, three distinct clusters or sublineages were observed. The first two sublineages (I and II) possessed the np 16111T mutation, with the first (I) also having the np 16192T mutation and the second (II) having the np 16265G mutation which arose in Eskimoan populations, whereas the third sublineage (III) lacked the 16111T mutation altogether. Sublineage III sequences appeared in haplogroup A mtDNAs from east Asian populations and the Evenks and probably represent the ancestral state for this mtDNA lineage. Among Paleoasiatic speakers, the Koryaks and Itel'men had haplogroup A CR sequences from both sublineages I and III, whereas those of the Chukchi belonged to sublineages I and II.

The previously identified CR sequence sublineages in haplogroup C were also ob-

Fig. 6. An NJ tree of CR sequences from northeast Siberian and east Asian populations constructed from genetic distances estimated with the Kimura two-parameter model in DNADIST (Felsenstein, 1994). The CR sequences occurring in the Koryaks and Itel'men were indicated as Kamchatkan (KAM) samples, and those occurring in the Chukchi and Siberian Eskimos (Starikovskaya et al., submitted) were indicated as Chukotkan (CHU) haplotypes, with the numbers of each corresponding to the CR sequence numbers in Table 8 from this study and those described in Starikovskaya et al. (1998). Otherwise, the CR sequences occurring in native Siberians or east Asians were indicated as the CR sequences enumerated in Table 5 of Torroni et al. (1993b). AF62 is the African CR sequence used as an outgroup in this tree, although the same overall branching structure was maintained when no outgroups were used. The Roman numerals specify distinct clusters or sublineages of the particular haplogroup in which they occur, while all haplogroups are indicated by boxed capital letters.

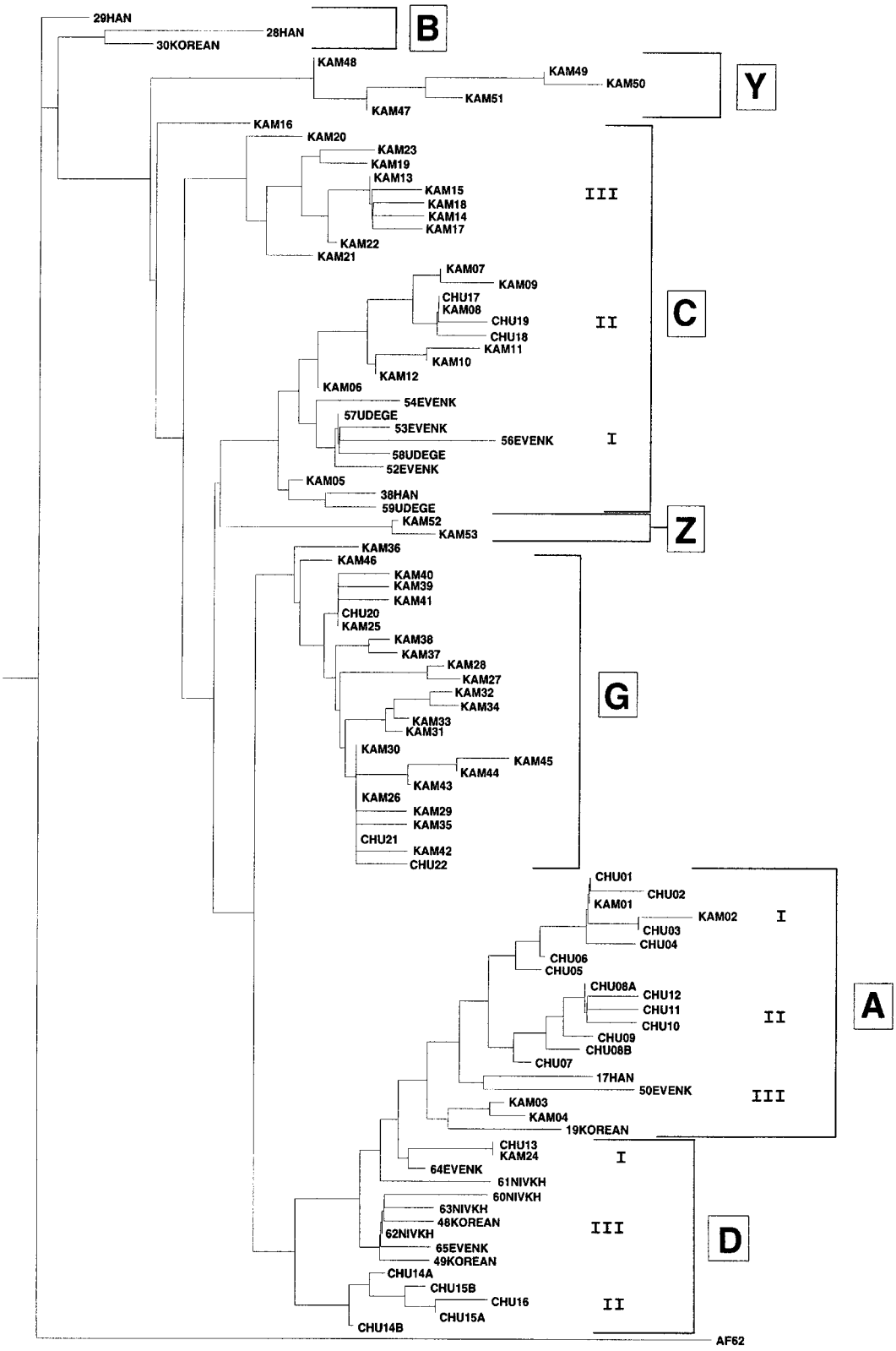
29HAN
28HAN
30KOREAN
B
KAM48
KAM49
KAM50
KAM51
KAM47
Y
KAM16
KAM20
KAM23
KAM19
KAM13
KAM15
KAM18
KAM14
KAM17
KAM22
KAM21
III
KAM07
KAM09
CHU17
KAM08
CHU19
CHU18
KAM11
KAM10
KAM12
KAM06
II
C
54EVENK
57UDEGE
53EVENK
56EVENK
58UDEGE
52EVENK
KAM05
38HAN
59UDEGE
KAM52
KAM53
I
Z
KAM36
KAM46
KAM40
KAM39
KAM41
CHU20
KAM25
KAM38
KAM37
KAM28
KAM27
KAM32
KAM34
KAM33
KAM31
KAM30
KAM45
KAM44
KAM43
KAM26
KAM29
KAM35
CHU21
KAM42
CHU22
G
CHU01
CHU02
KAM01
KAM02
CHU03
CHU04
CHU06
CHU05
I
CHU08A
CHU12
CHU11
CHU10
CHU09
CHU08B
CHU07
II
A
17HAN
50EVENK
KAM03
KAM04
19KOREAN
III
CHU13
KAM24
64EVENK
61NIVKH
60NIVKH
63NIVKH
48KOREAN
62NIVKH
65EVENK
49KOREAN
I
III
D
CHU14A
CHU15B
CHU16
CHU15A
CHU14B
II
AF62

Fig. 6.

served in this NJ tree. Two of them more closely resembled the haplogroup C mtDNAs present in Native American populations by having the 16223T-16298C-16327T motif characteristic of this mtDNA lineage. The one defined only by these mutations (I) occurred predominantly in east Asian and Amur River populations as well as the Evenks, and the other (II), having the 16124C and 16318T mutations, occurred only in Paleoasiatic groups. The third sublineage (III) had the 16093C-16189C-16261T-16288C-16298C motif and appeared only in northern Paleoasiatic-speaking groups.

Similarly, haplogroup D had three sublineages within it. The first (I), representing SIB40, occurred at very low frequencies among Paleoasiatic groups and Siberian Eskimos and was the only haplogroup D mtDNA in the Koryaks. The second sublineage (II) occurred exclusively among Chukotkan populations, as no similar types were seen among Native American groups with high frequencies of haplogroup D (Ward et al., 1991, 1993; Shields et al., 1993). The remaining mtDNAs from east Asian populations, the Evenks, and the Nivkhs formed a sublineage (III) which had a sequence motif most similar to haplogroup D haplotypes in Native American populations (16223T–16362C) and thus probably represented the ancestral state for this mtDNA lineage in Asia and the Americas.

Comprised of only Koryak, Itel'men, and Chukchi mtDNAs, haplogroup G also showed some degree of substructure. However, its branches were not quite as clearly defined as the sublineages in haplogroups A, C, and D. Instead, it exhibited small clusters of related CR sequences which appeared mostly in either the Chukchi, Koryaks, or Itel'men, which may reflect its relatively more recent origin in eastern Siberia relative to the other three haplogroups. By contrast, haplogroups B, Y, and Z formed small unbranched clusters, probably due to the limited number of mtDNAs analyzed for each mtDNA lineage in east Asian and Siberian populations. Notably, only east Asian mtDNAs were found within haplogroup B, while Koryak and Itel'men mtDNAs constituted all of the CR sequences present in the other two haplogroups.

Genetic links between eastern Siberian and east Asian populations

Because the distribution of haplogroups G and Y in northeast Asia pointed to the Sea of Okhotsk region as a possible source area for these mtDNA lineages, we examined the mtDNA variation in populations from this geographic region to determine their genetic affinities with Paleoasiatic groups. In particular, we were interested in assessing the relatedness of the Ainu, Japanese, and Koreans to the Koryaks and Itel'men since these east Asian groups had also been shown to have haplogroup G and/or Y mtDNAs (Horai et al., 1984; Ballinger et al., 1992; Harihara et al., 1992). To do this, we compared the CR sequence data from the Ainu, Japanese, and Koreans (Torroni et al., 1993b; Horai et al., 1996) with those from Siberian Eskimos, Paleoasiatic-speaking groups, Amur River populations (Nivkhs and Udegeys), and the Evenks. Although the sequences of Horai et al. (1996) lacked sequence information for the region between np 16000 and 16048 in which the np 16017C mutation from haplogroup G occurs, all of the other phylogenetically important nucleotide polymorphisms for haplogroups A–D, G, Y, and Z were contained within the region encompassed by their sequences (np 16048–16530). Hence, inferences about the relationships of the mtDNAs from these populations with Siberian groups was possible.

The resulting NJ tree revealed a number of interesting associations between these populations (Fig. 7). The first notable finding was that most of the haplogroups defined in this NJ tree occurred in both Paleoasiatic and east Asian populations. Although not previously classified as belonging to haplogroup A, due to the HaeIII np 663 site gain not being clearly identified in the earlier RFLP study of these samples (Horai and Matsunaga, 1986), a small number of Ainu, Japanese, and Korean sequences from Horai et al. (1996) clearly fell into this mtDNA lineage (cluster C6 [Horai et al., 1996]). All of these sequences lacked the np 16111T mutation and hence were part of the Asian/Siberian sublineage III of this haplogroup (Fig. 6). In addition, the Koryaks and the Ainu shared CR sequence #03 from this

study, and the Itel'men had the closely related CR sequence #04 from this haplogroup. By contrast, the Japanese and Koreans exhibited a set of related CR sequences from haplogroup A which were distinctive from those in Kamchatkan groups by having the np 16187T mutation and also lacking the 16362C transition.

Second, unlike Siberian populations, all three of the east Asian groups analyzed by Horai et al. (1996) had haplogroup B mtDNAs (cluster C2 [Horai et al., 1996]). The Ainu had a very low frequency of these haplotypes, as seen in an earlier RFLP study (Harihara et al., 1988), and those present in this population were very similar to types detected in the Japanese and Koreans, who were also previously noted to possess deletion haplotypes (Horai and Matsunaga et al., 1986; Ballinger et al., 1992). Based on this distribution, it appears that the Ainu acquired deletion haplotypes through gene flow with Japanese populations rather than having them as part of their ancestral gene pool. In addition, haplogroup B mtDNAs separated into two distinct subbranches rather than remain in a single cluster, as seen in Horai et al. (1996). This result is in concordance with findings in other studies of Asian mtDNA variation (Ballinger et al., 1992; Schurr, Starikovskaya et al., unpublished) which argue against this mtDNA lineage being a monophyletic group.

Third, all three east Asian groups (Ainu, Japanese, Koreans) had haplogroup C mtDNAs at low frequencies (cluster C14 [Horai et al., 1996]). This mtDNA lineage was previously observed at low frequencies in the Japanese (morph-9 [Horai et al., 1984]) but had not been detected in the Koreans (Ballinger et al., 1992) or the Ainu (Harihara et al., 1988). The majority of these CR sequences were located in sublineage I defined by the 16223T-16298C-16327T motif (Fig. 6), with all of them belonging to the Japanese and Koreans. In contrast, a minority of these CR sequences showed affinities with sublineage III in Paleoasiatic groups, including those present in the Ainu, although only the Koreans had closely related mtDNAs. This distribution again suggested an eastern Siberian/east Asian source for sublineage I mtDNAs within this haplogroup and perhaps an east Asian source for sublineage III mtDNAs.

Another novel finding was that haplogroup D mtDNAs appeared to be very common in the Ainu, Japanese, and Koreans. In fact, a number of different clusters of CR sequences in these populations (clusters C5, C7, C8, C9, C10, and C11 [Horai et al., 1996]) had the 16223T–16362C motif which characterizes this haplogroup. In general, the east Asian CR sequences from these clusters were interspersed among similar types from eastern Siberian groups, including the neighboring Nivkhs (sublineage II; Fig. 6), with the only exceptions being the cluster C11 mtDNAs, which, in having the np 16189C mutation, formed a separate subbranch. Similarly, the Chukchi and Eskimo CR sequences from SIB48–53 also formed a separate branch (sublineage III; Fig. 6), as did that of SIB40 (sublineage I; Fig. 6), which was positioned closest to haplogroup A, due to having several polymorphisms in common with this mtDNA lineage. This branching pattern further illustrated the significant diversity of CR sequences within the haplogroups present in east Asian and Siberian populations and indicated that many of the CR clusters seen in Figure 7 are not monophyletic groupings equivalent to haplogroups, as suggested by Horai et al. (1996).

Similarly, CR sequences from haplogroup G were apparently present in all three east Asian populations, a result which was consistent with previous RFLP analyses of Korean (Ballinger et al., 1992) and Japanese (Horai et al., 1984; Harihara et al., 1992) populations. The Ainu showed the highest frequency of these types (cluster C16 [Horai et al., 1996]), and nearly all of them clustered among putatively similar mtDNAs from Paleoasiatic populations, implying a common genetic origin. On the other hand, all of the Japanese and Korean CR sequences in this haplogroup cluster together separate from the Siberian/Ainu branch, implying that they might represent a divergent sublineage of this mtDNA lineage which arose in their common ancestral population.

Concerning haplogroup Y, previous RFLP studies of mtDNA variation (Horai et al.,

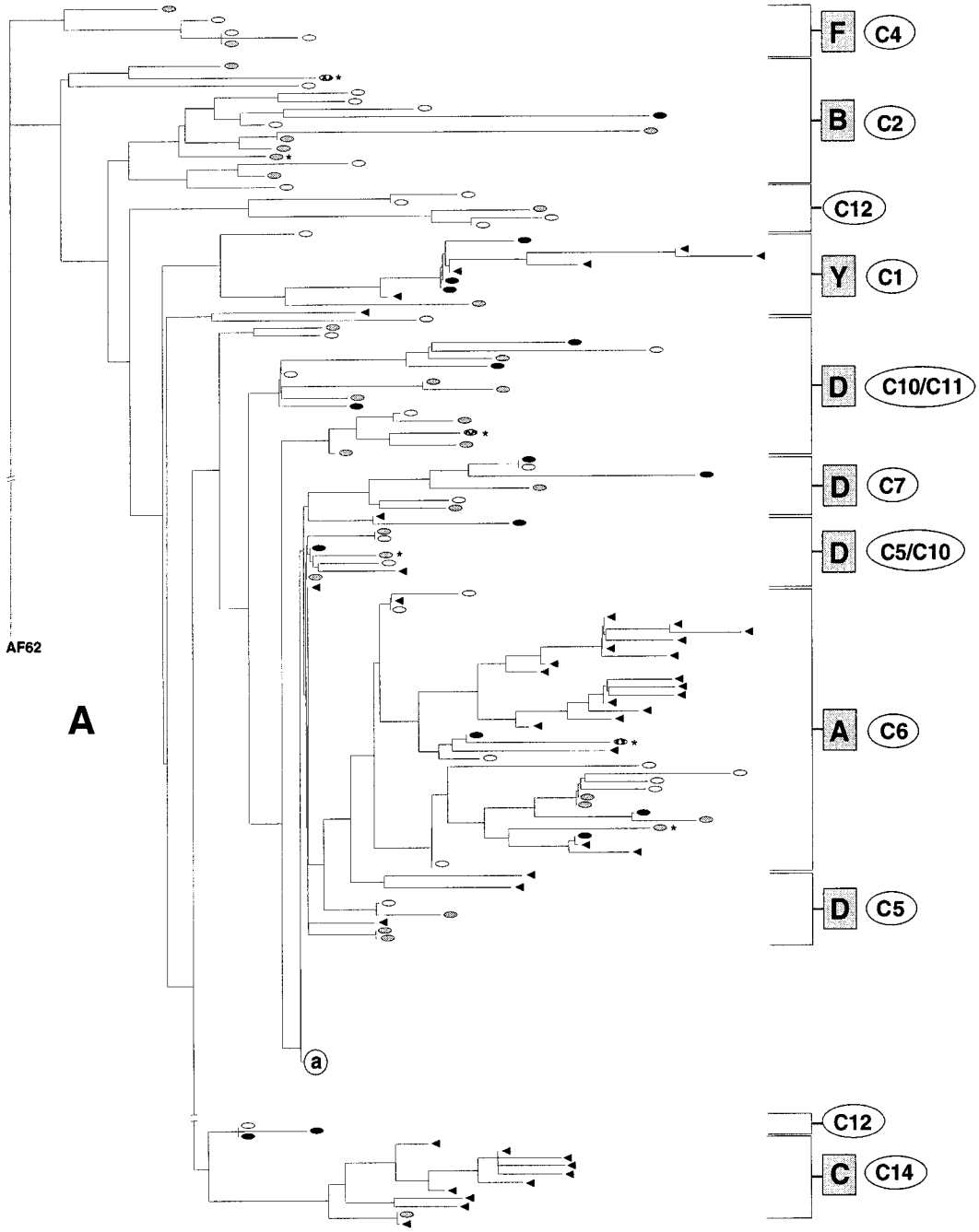

Fig. 7. An NJ tree of Siberian and east Asian CR sequences based on genetic distances estimated with the Kimura two-parameter model in DNADIST (Felsenstein, 1994). All Siberian individuals analyzed in Starikovskaya et al. (submitted) and this study are indicated by branches with black triangles at their terminal tips, while those from Ainu, Korean, and Japanese individuals analyzed in Horai et al. (1996) are indicated by ellipses, with the population affiliation specified in the key. The Korean and Taiwanese Han samples analyzed by Torroni et al. (1993b) are also indicated by ellipses but are denoted with asterisks to distinguish them from those analyzed in Horai et al. (1996). The haplogroups to which the sequences belonged or were assigned based on their CR sequence motif and position in this NJ tree are indicated by the capital letters. The capital letters in

Fig. 7. (Continued) shaded boxes correspond to the haplogroups of Torroni et al. (1993a,b, 1994c) and this study, whereas the groupings specified with a capital C followed by a number correspond to the CR sequence clusters identified in Horai et al. (1996). When there was an exact correlation between these cluster designations, both were positioned by the bracket which encloses the related CR sequences. No Siberian population showed CR sequences which belonged to clusters C7, C9, C11, C13, C15, and C17. Cluster C3 did not appear in this tree because its sequences did not appear in the Ainu, Japanese, or Koreans, and C18 is absent because only one Korean from Horai et al. (1996) had a sequence belonging to it. The circled a in both panels of the figure indicates where the two portions of the NJ tree are connected.

1984; Horai and Matsunaga, 1986) were not able to show the presence of these haplotypes in the Japanese. However, CR sequence analysis confirmed the presence of haplogroup Y mtDNAs in the Japanese and Koreans at very low frequencies as well as in 19.6% of the Ainu from Hokkaido (cluster C1 [Horai et al., 1996]). These findings suggested that Paleoasiatic and Ainu populations had haplogroup Y mtDNAs as part of their ancestral genetic makeup or possibly that the Itel'men and Koryaks acquired some haplogroup Y mtDNAs through contact with Ainu populations (Near Kurilers) who occupied the southern tip of the Kamchatkan population in prehistoric times. Furthermore, the occurrence of haplogroup Y mtDNAs in the Japanese and Koreans implied that these populations obtained them through gene flow with the Ainu or other Siberian groups having these mtDNAs.

Very few if any east Asian CR sequences clustered with haplogroup Z mtDNAs from Kamchatkan populations (cluster C14 [Horai et al., 1996]). All of the Japanese and Korean CR sequences which clustered near these types had at least two of the defining mutations of this mtDNA lineage (16185T and 16260T and/or 16298C) but lacked two others (16129A, 16224C) which were present in all comparable Kamchatkan mtDNAs. This finding suggested that, if part of haplogroup Z, these mtDNAs were distantly related to those appearing in the Koryaks and Itel'men. These differences, along with the presence in Mongolians (Kolman et al., 1996) and Evens (Table 1) of haplogroup Z CR sequences identical to those in Kamchatkan populations, probably means that this mtDNA lineage did not evolve in the Sea of Okhotsk/Amur River region.

There were also several other clusters of CR sequences from the east Asian populations which did not appear in Siberian populations. At least one of these was haplogroup F mtDNAs, since this mtDNA lineage was known to be present in the Japanese (Horai et al., 1984; Harihara et al., 1992), Koreans (Ballinger et al., 1992), and Ainu (Harihara et al., 1992). Based on preliminary CR sequence data for Southeast Asians (Schurr et al., unpublished data), this branch can tentatively be identified as cluster C4 in Horai et al. (1996). However, in Figure 7, these mtDNAs were split into two clusters, with most CR sequences belonging to the large cluster located between cluster C17 of Horai et al. (1996) and haplogroup D.

Due to the lack of RFLP data for the remaining east Asian CR sequences, the exact lineal affiliations of the other two main clusters (C15 and C17 [Horai et al., 1996]) remains unknown, although both have very distinctive CR sequence motifs. C15 and C17 mtDNAs appeared in the Japanese and Koreans but occurred at the highest frequency in the Ainu and the Ryukyuans of Okinawa, the two aboriginal populations of the Japanese archipelago. Because these clusters were not present in eastern Siberian populations and occurred at low frequencies in Koreans and Japanese, they must have evolved in ancestral populations of the Ainu and Ryukyuans which expanded into these islands before the ancestors of the modern Japanese and Koreans arrived. This interpretation is generally consistent with the hybridization hypothesis for Japanese origins (Hanihara, 1991) which proposes that Jomon peoples originated in Southeast Asia and came to the Japanese islands $>$12,000 YBP, where they gave rise to the Ainu and Ryukyuans, while the progenitors of the Japanese and Koreans, the Yayoi people, emigrated from the Korean peninsula some 2,300 YBP and replaced or absorbed these aboriginal groups.

The relationships shown in Figure 7 were completely consistent with the nucleotide diversity estimates calculated for Paleoasiatic and east Asian populations (data not shown). These estimates, along with the associated NJ tree (Fig. 8), clearly showed the close genetic affinities of contemporary Korean and Japanese populations as well as the genetic similarity of the Ainu to both groups, with the latter association probably being attributable to recent admixture between the Ainu and Japanese. Moreover, there was the large split between Paleoasiatic-speaking populations, with the Koryaks and Itel'men showing much closer genetic ties to the Ainu and the Chukchi being more closely linked with the Siberian Eskimos and Northwest Coast Amerindian populations. These associations were consistent

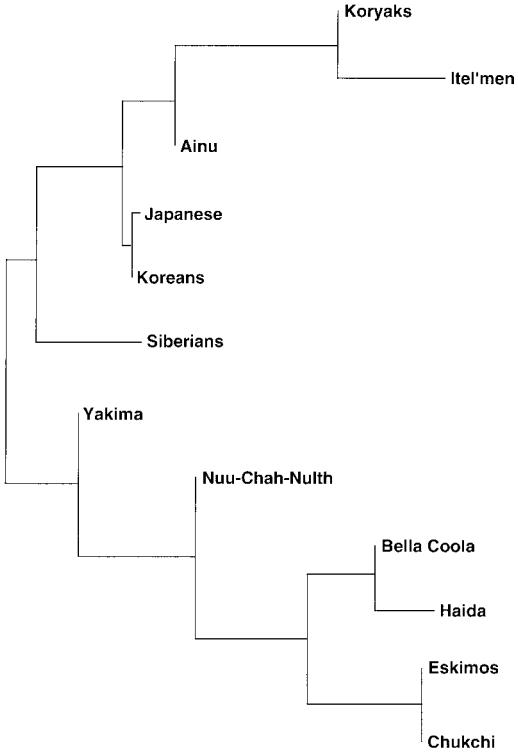

Fig. 8. An NJ tree based on genetic distances generated from pairwise nucleotide diversity estimates for Siberian and east Asian populations. Branch lengths are proportionate to the relative distances between populations. The sources and designations for these populations are specified in Methods.

with previous phylogenetic evidence and statistical estimates of genetic identity by descent which implied the considerable divergence of Paleoasiatic-speaking populations (Tables 6, 7) as well as with similar comparisons of Native American and east Asian CR sequence diversity (Shields et al., 1993; Horai et al., 1996).

DISCUSSION

Genetic history of Kamchatkan populations

One of the most significant findings of this study was the genetic discontinuity between Paleoasiatic-speaking populations of Chukotka and Kamchatka. Several different measures of haplotypic diversity showed that the Koryak and Itel'men populations were genetically very similar to one another but quite distinct from the Chukchi, who are linguistically related to the Koryaks. The Kamchatkan groups were also quite divergent from those which evolved from the ancient Beringian gene pool, such as the Eskimo-Aleuts and Na-Dené Indians, suggesting that their ancestral populations replaced the survivors of the Bering land bridge in this region during the Neolithic period. The only links to these earlier populations were the presence in the Koryaks of haplotype SIB41 mtDNAs from haplogroup A having both the 16111T and np 16192T mutations and of haplotype SIB40 from haplogroup D. However, the recent acquisition of these haplotypes through gene flow with the neighboring Chukchi, in whom these types of mtDNAs are more common, cannot be excluded.

While more closely related to each other than to any other Siberian population, the Koryaks and Itel'men showed significant differences between them. The CR sequence data did show that Koryaks and Itel'men shared the putative founding mtDNAs of haplogroups C, G, Y, and Z, suggesting they might have originated from a common ancestral population in the Okhotsk Sea region. However, they also exhibited significant differences in haplogroup frequencies and haplotype distributions, with nearly all unique CR sequences occurring in one population or the other. These results support other linguistic and culture evidence that the Itel'men and Koryak populations arose from temporally distinct expansions into the Kamchatka peninsula, with the ancestral Itel'men being the first to enter this region during the Siberian Neolithic (Arutiunov, 1988; Dikov 1990, 1994). These results further reveal an inconsistency between the linguistic affiliations and genetic profiles of Paleoasiatic speakers, a phenomenon which appears to be common for most indigenous populations of Siberia (Szathmary, 1981; Sukernik, 1992).

Moreover, both RFLP and CR sequence data further revealed the considerable differentiation of Koryak subgroups living in Kamchatka. Although sharing several putative founding mtDNAs, the Aluitor, Karagin, and Palan Koryaks exhibited significant differences in haplogroup frequencies and haplotype distribution. These results were consistent with both ethnographic and linguistic evidence for dialectic subdivisions of the

Koryak population, as summarized by Vdovin (1973). In addition, the Aluitor and Karagin Koryaks appeared to be more genetically similar to each other than either was to the Palan Koryaks, as expected from their closer linguistic association, although differences in the distribution of the two major CR sublineages from haplogroup C (I and II) separated the Aluitor Koryaks from the other two subgroups. This apparent discrepancy is probably attributable to considerable admixture between Itel'men populations and the Karagin and Palan Koryaks, as marital exchanges between the Itel'men and Koryaks living along the Tigil' River have been documented since the seventeenth century (Krasheninnikov, 1972) and records of the Russian Orthodox Church from the village of Ivashka dating between ~1850 and 1930 (Sukernik and Schurr, unpublished data) indicate that the remnants of the eastern Itel'men fused with the adjacent Karagin Koryaks during the second half of the nineteenth century.

The high statistical significance of the differences in haplotype distribution in Kamchatkan populations and between the Koryak subgroups was also notable for other reasons. The closing of traditional settlements throughout the peninsula during the Soviet period led to the consolidation of widely separate settlements, which in turn probably caused the mixing of different Koryak subgroups. The most direct evidence of the effects of population amalgamation may be that Reindeer and Maritime Koryaks are largely genetically indistinguishable from one another despite the fact that Reindeer Koryaks developed out of Okhotsk Koryak populations. In addition, warfare, epidemics, and intermarriage with nonnatives in the last 300 years have had a profound effect on their population size and composition. Nevertheless, our results suggest that remnants of the former dialectical and territorial subdivisions of Koryaks have persisted into modern times despite the enormous demographic impact of Russian colonization.

The mtDNA data were also consistent with the archeological evidence from northeastern Siberia. Since 14,000 YBP, there has been a series of population expansions along the Asiatic coastline of the Bering Sea by cultures having different subsistence strategies and lithic technologies. The Early Ushki site in the Kamchatka peninsula (14,300 YBP) was occupied by a culture based on big game hunting and salmon fishing which employed bifacial projectile points similar to the Late Paleolithic Paleoindian stemmed point industry of western North America (Arutiunov and Sergeev, 1990a,b; Dikov, 1990, 1994). However, by 10,860 YBP, the Late Ushki culture differed markedly from the earlier one in terms of the size and form of its dwellings, the use of bifacially retouched leaf-shaped microblades, and the first appearance of stone lip labret ornamentation. These innovations in the early Holocene, along with the development of nontoggling, multibarbed harpoon technologies for hunting sea mammals which were widely employed in the Sea of Okhotsk and Bering Sea regions, demarcate the beginnings of the proto-Eskimo-Aleut cultural tradition in northeast Siberia (Arutiunov and Sergeev, 1990a,b; Dikov, 1990, 1994).

During the mid-Holocene climatic optimum, at approximately 6,000–4,000 YBP, there was substantial population growth in the littoral area of the Okhotsk Sea region associated with the spread of continental cultures of reindeer hunters from the Lena and Kolyma River basins (Mochanov, 1962; Fedoseeva, 1968; Simchenko, 1976). As suggested by Vasilievskiy (1971), the expansion of these continental tribes into the northern Okhotsk Sea region apparently gave rise to the ancestral Koryak and Itel'men populations, whereas movement from the lower Amur River–Sakhalin region appeared to have played a supplementary role in their origins. Furthermore, the expansion of Neolithic "Southern Okhotsk" cultures into northern Japan from the lower Amur River region might have substantially contributed to the origin of the Ainu, who, until very recently, occupied Sakhalin Island, the Kurile Islands, and the southern tip of the Kamchatka peninsula (Vasilievskiy, 1971; Krasheninnikov, 1972; Arutiunov, 1988). This scenario, based primarily on archeological data, suggests that the genetic profiles of the Koryaks and Itel'men should be distinctive from those of the Nivkhs and Ainu, an

TABLE 10. Sequence divergence of mtDNA haplogroups in Siberia and the Americas[1]

Haplogroup	Geographic region	n	N	Sequence divergence (%)	Divergence time (YBP)
A	Siberia	10	119	0.0280	12,714–9,645
	America	46	189	0.0789	35,550–26,969
B	America	30	99	0.0391	17,773–13,483
C	Siberia	14	123	0.0433	19,686–14,934
	America	31	77	0.1223	54,009–40,972
D	Siberia	13	47	0.1115	50,664–38,434
	America	16	62	0.0565	25,682–19,483
G	Siberia	11	106	0.0239	10,855–8,234
Y	Siberia	7	58	0.0138	6,273–4,759
Z	Siberia	4	12	0.0209	9,495–7,203

[1] n, number of haplotypes for each haplogroup; N, number of individual mtDNAs for each haplogroup. The sequence divergence estimates were weighted by the number of individuals within each haplogroup, and divergence times were calculated using a mtDNA evolutionary rate of 2.2–2.9% per MYR (Torroni et al., 1994a).

interpretation which is supported by the mtDNA data.

On a broader scale, the expansion of Paleoasiatic-speaking peoples into northeast Asia led to the near total replacement of the ancient Bering Sea cultures in Kamchatka, with different varieties of the ancient Koryak culture diffusing extensively along the Okhotsk Sea and coastline of the northwestern Pacific (Vasilievskiy, 1971; Arutiunov and Sergeev, 1990a,b; Dikov, 1994). However, the mtDNA data indicate that, while absorbing elements of the Eskimo-Aleut culture during their expansion, ancestral Koryak and Itel'men groups did not extensively incorporate members of these maritime tribes (Torroni et al., 1993b; Shields et al., 1993; Starikovskaya et al., 1998; this study). This pattern was also seen in the Y-chromosome data for the same populations (Lell et al., 1997a,b), which showed strong links between Native American and Chukotkan populations and their distinctiveness from other northeast Asian groups. Thus, both the genetic and archeological data indicate that multiple population and/or cultural expansions have taken place in the Okhotsk Sea and Bering Sea region over the last 10,000 years, with more recently evolved genotypes and cultural traditions from northeast Asia overlapping and/or replacing more ancient ones.

Genetic discontinuity at the north Pacific Rim

The analysis of mtDNA variation within and among Paleoasiatic speakers (Chukchi, Koryaks, and Itel'men), Eskimos, Na-Dene Indians, and Amerindian tribes of the Pacific Northwest have shown that these groups are quite divergent from one another (Shields et al., 1993; Torroni et al., 1993a,b; Starikovskaya et al., 1998; this study). Although additional mtDNA lineages are present in all Siberian populations except for the Siberian Eskimos, the common ancestry of Siberian and Native American groups is evidenced by the ubiquitous presence of haplogroups A, C, and D in these populations (Table 4). However, aside from the putative founding haplotypes for haplogroups A (SIB41/AM09), C (SIB26/AM43), and D (SIB13/AM88), these populations share no other haplotypes, with the remaining mtDNAs from these haplogroups being largely population- or region-specific. Along with the ML estimates for haplogroups A, C, and D in both regions (Table 10), these findings imply the considerable antiquity of the primary mtDNA lineages occurring in both Siberia and the New World as well as their extensive divergence from each other since being isolated in each continental region over 20,000 years ago.

While evidence for the antiquity of the initial colonization of the New World is rapidly accumulating (e.g., Bonatto and Salzano, 1997a), the population dynamics in northeast Asia subsequent to the last glacial maximum (~18,000 YBP) are of more importance for determining the origins of Paleoasiatic speakers of Chukotka and Kamchatka and their affinities with other Siberian populations. In this regard, certain of the major populational events occurring during this period can be associated with specific mtDNA

polymorphisms present in haplogroup A mtDNAs. To begin with, the 16111T mutation in haplogroup A mtDNAs delineates the emergence of ancestral Paleoindian populations and their dispersal in the New World. Later, after the initial occupation of the New World, ancient Beringian populations apparently became isolated from ancestral Paleoindian groups, during which time the large north Pacific Rim sublineage defined by the 16192T mutation arose among the ancestral populations for the Chukchi, Siberian and Alaskan Eskimos, and Na-Dené Indians. In addition, a number of population- or region-specific haplotypes in each of these two sublineages of haplogroup A arose in the Beringian groups independent of those occurring in Paleoindian populations (Ward et al., 1991, 1993; Torroni et al., 1992, 1993b; Shields et al., 1993; Starikovskaya et al., 1998). A similar pattern of diversity was also observed for haplogroup D mtDNAs in Chukotkan populations, which differed from comparable Native American mtDNAs by several unique mutations (Starikovskaya et al., in press). These population-specific mtDNA sublineages probably reflect the isolation and reemergence of remnant populations occupying biogeographic refugia in Beringia and southern Alaska which existed until the end of the last glacial maximum (Rogers et al., 1991).

The pattern and timing of the expansions out of Beringia are also mirrored by the different divergence values for haplogroup A in Siberia and the Americas (Table 10). The estimated sequence divergence for this haplogroup in Siberia was 0.028%, a value considerably less than that for the Americas, 0.079%. These values give correspondingly different divergence times for Siberia (13,000–10,000 YBP) and the Americas (36,000–27,000 YBP). This apparent discrepancy is largely attributable to almost exclusively Chukotkan haplotypes being present in the haplogroup A estimate for Siberia. In fact, our estimates of the genetic divergence of haplogroup A in Siberian and Native American populations, one for Chukotkan groups (0.029%, 12,727–9,655 YBP), another for Na-Dené Indians (0.021%, 9,545–7,241 YBP [Torroni et al., 1992]), and a third for Amerindians (0.079%, 35,909–27,241 YBP), clearly show the extent of diversity which has developed in them, not just within the haplogroup itself. This interpretation is supported by the fact that only the founding haplotype (SIB41/AM01) of this haplogroup is shared amongst them, and all other haplotypes are unique to each set of populations (Torroni et al. 1992, 1993a,b, 1994a,b; Starikovskaya et al., 1998). Thus, while these divergence estimates do not give exact times for the ages of specific ethnic groups, they do provide a temporal framework in which to view the emergence of the ancestral populations for the three major Native American linguistic divisions.

Given this pattern of genetic divergence in northeast Asia and the New World, it was not surprising that there was a striking discontinuity in haplotypic diversity between Kamchatkan and Native American populations. Although having a number of haplogroup A, C, and D haplotypes, the Koryaks and Itel'men were not closely genetically related to Native American groups and actually shared only SIB41 (AM01) from haplogroup A and SIB26 (AM43) and SIB45 (AM32) from haplogroup C. The NJ tree of CR sequences (Fig. 6) strongly confirmed the pattern seen in the RFLP haplotype data, with the Kamchatkan CR sequences having the strongest affinities with those in Native Americans belonging only to haplogroups A and C. However, the Koryak haplogroup A mtDNAs (SIB41) having the 16111T mutation also possessed the 16192T mutation, indicating that they were not directly linked to those in Amerindian groups. The other haplogroup A CR sequences in the Koryaks and Itel'men lacked the 16111T mutation and in doing so more closely resembled mtDNAs present in east Asian and eastern Siberian groups, in whom they probably originated. Therefore, while Kamchatkan groups have haplogroup A mtDNAs and even the putative founding haplotype, SIB41 (AM01), none of them are closely related to comparable mtDNAs from Amerindian populations.

The same trend was observed in haplogroup C mtDNAs from Kamchatkan populations. The Siberian mtDNAs with CR sequences most akin to those in Native Americans (sublineage I) represented a mi-

nority of those present in the Koryaks (11.9%) and did not occur in the Itel'men at all, whereas they occurred at much greater frequency in other eastern Siberian (Evenks and Udegeys) and east Asian (Han Chinese and Koreans) populations (Torroni et al., 1993b). The remaining haplogroup C mtDNAs in Kamchatkan populations belonged to sublineage II, which occurred only in the Koryaks, and sublineage III, which occurred in both Koryaks and Itel'men, with sublineage II being related to haplotypes present in east Asian and eastern Siberian populations. Thus, while SIB26 haplotypes comprised 23.2% of all Kamchatkan mtDNAs from haplogroup C, very few of them resembled comparable mtDNAs in Native American groups at the CR sequence level. Therefore, the majority of the Koryak and Itel'men mtDNAs are not the same as the founding haplotype in New World populations and instead must have arisen after the colonization of the New World.

Our analyses have also shown that haplogroup B mtDNAs are absent in the Koryaks and Itel'men as well as the Chukchi and Siberian Eskimos (Torroni et al., 1993b; Starikovskaya et al., in press; this study). Their absence in the Koryaks and Itel'men was an important finding because the Kamchatka peninsula was contiguous with the rest of Beringia during the last glacial maximum (Fladmark, 1979; Hopkins, 1979; Hoeffecker et al., 1993) and could have been part of the route that the immigrant population(s) carrying haplogroup B took while moving across the Bering Strait from Asia into the Americas. Thus, it appears that haplogroup B was never part of the ancestral gene pool for Paleoasiatic-speaking populations and that these populations played no role in the dispersal of this mtDNA lineage into the New World.

In contrast, the Ainu, with whom the Itel'men reportedly had considerable contact in historic times (Krasheninnikov, 1972), exhibited haplogroup B mtDNAs, although at low frequencies (2.0%) (Harihara et al., 1992). Along with the higher frequencies of haplogroup B in the modern Japanese and Koreans (~10–16%) (Horai et al., 1984; Ballinger et al., 1992; Horai et al., 1996) this finding may reflect the relatively recent expansion of this mtDNA lineage into the Sea of Okhotsk region. These data also suggest that populations bearing haplogroup B mtDNAs could have originated in east Asia and moved across Beringia via a coastal route.

Several lines of evidence support the hypothesis of a separate migration of peoples carrying haplogroup B mtDNAs through Beringia to the New World. First, the absence of haplogroup B mtDNAs in central and eastern Siberian populations which share founding haplotypes from haplogroups A, C, and D with Native American groups may imply that these haplotypes were not present in the original progenitors of New World populations. Second, the virtual absence of haplogroup B mtDNAs in modern Eskimo, Aleut, and northern Na-Dené Indian populations, which represent more recent demic expansions into North America, implies that haplogroup B mtDNAs were not present in the Beringian region after 10,000 YBP, when these populations were founded. Third, the ML divergence estimate for haplogroup B of 17,000–13,000 YBP is considerably smaller than that of haplogroups A, C, and D in the New World (Table 10). These data imply that haplogroup B mtDNAs arrived in the Americas after the initial Paleoindian migration brought haplogroups A, C, and D to the New World but before the Beringian expansion(s) which gave rise to the ancestral Eskimo-Aleut and Na-Dené Indian populations into the arctic and subarctic regions. This later expansion of Beringian groups into the New World might also explain the low frequencies of haplogroup B mtDNAs in most northern North American Indian populations (Ward et al., 1991, 1993; Torroni et al., 1992, 1993a; Lorenz and Smith, 1994) relative to those inhabiting regions farther south.

Alternatively, haplogroups A–D may have been brought together during the initial colonization of the New World. All four haplogroups are observed in most modern and ancient Native American populations (Schurr et al., 1990; Ward et al., 1991, 1993; Torroni et al., 1992, 1993a, 1994a,b; Ginther et al., 1993; Horai et al., 1993; Santos et al., 1994; Stone and Stoneking, 1994; Batista et al., 1995; Kolman et al., 1996; Lorenz and

Smith, 1994, 1996; Merriwether et al., 1995). The distribution and age of haplogroup B in Asia also suggest that this mtDNA lineage evolved in and spread from the region encompassing Mongolia, Tibet, the northern Himalayas, and southern Siberia by at least 30,000–24,000 YBP (Ballinger et al., 1992; Lum et al., 1994; Horai et al., 1996), implying it could have been present in the ancestral Siberian groups which first moved into the New World. Such dates are consistent with recent estimates of CR sequence diversity within haplogroup B in Native Americans, which indicate that this mtDNA lineage arrived in the New World by 30,000–25,000 YBP (Bonatto and Salzano, 1997b). These competing interpretations clearly indicate that further research is required to delineate the origin and dispersal of haplogroup B in the Americas.

On the other hand, haplogroups G, Y, and Z have not been observed among Native American groups analyzed by high resolution RFLP analysis (Schurr et al., 1990; Torroni et al., 1992, 1993a, 1994b). Although "Other" haplotypes have also been detected in both ancient and modern Native American populations by partial haplotype analysis and CR sequencing (Bailliet et al., 1994; Hauswirth et al., 1994; Stone and Stoneking, 1994; Merriwether et al. 1995; Lorenz and Smith, 1996; Ribiero-Dos-Santos et al., 1996), the limited data for these samples indicate that they also do not belong to haplogroups G, Y, and Z. Therefore, populations bearing these haplotypes must have spread in northeast Asia after the populating of the New World. This interpretation is consistent with their younger divergence times relative to those of haplogroups A–D in Siberia and the Americas (Table 10) and with haplogroup G being the oldest and most widespread of these mtDNA lineages in the eastern Siberia/east Asia region (Horai et al., 1984; Ballinger et al., 1992; Torroni et al., 1993b, 1994c; Starikovskaya et al., 1998; this study).

ACKNOWLEDGMENTS

The authors thank Lorri Griffin and the Clinical Research Center of the Emory University School of Medicine, supported by NIH M01-RR-00039, for their assistance in the processing of blood samples, and Drs. Andy Kogelnik and Sandro L. Bonatto for their assistance with the statistical analyses of the mtDNA data. We also thank the hospital staff and doctors in the villages of Ossora, Karaga, and Tymlat for their assistance with this project and the Koryak people from these villages for their participation in this research and hospitality during our field work in July–August 1993. Similarly, we extend our gratitude to the hospital staff in the villages of Voyampolka and Kovran for assistance with this project and to the Koryak and Itel'men people from these villages for their participation in this research during our field work in June 1996. This research was supported by grants from The Wenner-Gren Foundation for Anthropological Research (GR 5745), The National Science Foundation (NSF 9414900), The Leakey Foundation, and Sigma Xi to T.G.S.; from The Wenner-Gren Foundation for Anthropological Research (GR 5680, GR 5680-S, and GR 6077), INTAS (96–1766), and the Russian Fund for Basic Research (97–04–49816) to R.I.S.; from the J. Worley Brown Fellowship Fund and the National Institutes of Health (GM49615, NS21328, HL45572, AG10130, and AG13154) to D.C.W; and from the Emory Clinical Research Center (BR00039).

LITERATURE CITED

Anderson S, Bankier AT, Barrell BG, DeBruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Rose BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–465.

Antropova VV. 1964a. The Koryaks. In: Levin MG, Potapov LP, editors. The peoples of Siberia. Chicago: University of Chicago Press. p 851–875.

Antropova VV. 1964b. The Itel'mens. In: Levin MG, Potapov LP, editors. The peoples of Siberia. Chicago: University of Chicago Press. p 876–883.

Arutiunov SA. 1988. Koryaks and Itel'men: dwellers of the smoking coast. In: Fitzhugh WW, Crowell A, editors. Crossroads of continents. Baltimore: Smithsonian Institution Press. p 31–35.

Arutiunov SA, Sergeev DA. 1990a. Issues of the ethnic history of the Bering Sea. Chapter 1. Problems in the study of the ethnic history of the northwestern coastline of the Pacific coast. Soviet Anthropol Archeol 28:50–61.

Arutiunov SA, Sergeev DA. 1990b. Issues of the ethnic history of the Bering Sea. Chapter 8. The role of ancient Eskimo culture in the ethnic history of the northern Pacific basin. Soviet Anthropol Archeol 28: 62–77.

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO. 1994. Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55: 27–33.

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JH, Hassan K, Chen K-H, Wallace DC. 1992. Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152.

Batista O, Kolman CJ, Bermingham E. 1995. Mitochondrial DNA diversity in the Kuna Amerinds of Panamá. Hum Mol Genet 4:921–929.

Bogoras W. 1910. The Chukchee. Publication of the Jesup North Pacific Expedition, vol. 8, part 1. Leiden, NY: American Museum of Natural History. Reprinted New York: AMS Press, 1975.

Bonatto SL, Salzano FM. 1997a. A single and early origin for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc Natl Acad Sci U S A 94:1866–1871.

Bonatto SL, Salzano FM. 1997b. Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. Am J Hum Genet 61:1413–1423.

Cann RL, Brown WM, Wilson AC. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499.

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. Nature 325:31–36.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550–570.

Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 57:133–149.

Debets GF. 1951. Anthropological investigations in the Kamchatka Region. In: Transactions of the Institute of Ethnography, Russian Academy of Sciences, vol. 17. Moscow. p 3–263 (in Russian).

Derenko MV, Shields GF. 1998. Polymorphism in Region V of mitochondrial DNA in indigenous populations of northern Asia. Genetika 34:411–415.

Dikov NN. 1990. The origin of the aboriginal populations of America. Soviet Anthropol Archeol 28:12–29.

Dikov NN. 1994. The Paleolithic of Kamchatka and Chukotka and the problem of the peopling of America. In: Fitzhugh WW, Chausonnet V, editors. Anthropology in the Pacific Rim. Washington DC: Smithsonian Institution Press. p 87–95.

Easton RD, Merriwether DA, Crews DE, Ferrell RE. 1996. mtDNA variation in Yanomami: evidence for additional New World founding lineages. Am J Hum Genet 59:213–225.

Fedoseeva SA. 1968. Ancient cultures of the upper Vilui River. Moscow: Nauka (in Russian).

Felsenstein J. 1994. PHYLIP (Phylogeny Inference Package), version 3.572c. Seattle: University of Washington.

Fladmark KR. 1979. Routes: alternative migration corridors for early man in North America. Am Antiquity 44:55–69.

Forster P, Harding R, Torroni A, Bandelt HJ. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945.

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, Just J, Salzano FM, King MC. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplan JT, Jefferies AJ, eds. DNA fingerprinting state of the science. Basel: Berkhauser Verlag, Switzerland. p 211–219.

Gurvich IS. 1966. Ethnic history of north-east Siberia. Moscow (in Russian).

Hanihara K. 1991. Dual structure model for the population history of the Japanese. Japan Review 2:1–33.

Harihara S, Saitou N, Hirai M, Gojobori T, Park KS, Misawa S, Ellepola SB. 1988. Mitochondrial DNA polymorphism among five Asian populations. Am J Hum Genet 43:134–143.

Harihara S, Hirai M, Suutou Y, Shimizu K, Omoto K. 1992. Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. Hum Biol 64:161–166.

Hauswirth WW, Dickel CD, Rowold DJ, Hauswirth MA. 1994. Inter- and intrapopulation studies of ancient humans. Experientia 50:585–591.

Hoffecker JF, Powers WR, Goebel T. 1993. The colonization of Beringia and the peopling of the New World. Science 259:46–53.

Hopkins DM. 1979. Landscape and climate of Beringia during late Pleistocene and Holocene time. In: Laughlin WS, Harper AB, editors. The first Americans: origins, affinities, and adaptations. New York: Gustav Fischer. p 15–42.

Horai S, Hayasaka K. 1990. Intraspecific nucleotide sequence differences in the major non-coding region of the human mitochondrial DNA. Am J Hum Genet 46:828–842.

Horai S, Matsunaga E. 1986. Mitochondrial DNA polymorphism in Japanese. I. Analysis with restriction endonucleases of six base-pair recognition. Hum Genet 68:324–332.

Horai S, Gojobori T, Matsunaga E. 1984. Mitochondrial DNA polymorphism. II. Analysis with restriction endonucleases of four and five base-pair recognition. Hum Genet 72:105–117.

Horai S, Kondo R, Nakasawa-Hattori Y, Hayasaki S, Sonoda S, Tajima K. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Evol Biol 10:23–47.

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, Pan I-H. 1996. mtDNA polymorphism in east Asian populations, with special reference to the peopling of Japan. Am J Hum Genet 59:579–590.

Jochelson WR. 1908. The Koryak. The Jesup North Pacific Expedition, vol. 6. Leiden, NY: American Museum of Natural History. Reprinted New York: AMS Press, 1975.

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL. 1983. Radiation of human mitochondrial DNA types analysed by restriction endonuclease cleavage patterns. J Mol Evol 19:255–271.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. New York: Academic Press. p 21–132.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120.

Kolman CJ, Sambuughin N, Bermingham E. 1996. Mitochondrial DNA analysis of Mongolian populations and implications for the origins of New World founders. Genetics 142:1321–1334.

Krasheninnikov S. 1972 (1755). A description of the Kamchatka region, 1735–1741. Crownhart-Vaughn EAP, editor/translator. Portland, OR: Oregon Historical Society.

Krauss ME. 1988. Many tongues—ancient tales. In: Fitzhugh WW, Crowell A, editors. Crossroads of continents. Baltimore: Smithsonian Institution Press. p 145–150.

Krushanov AI. 1990. Itel'mens: history and culture. Leningrad: Nauka (in Russian).

Krushanov AI, editor. 1993. The history and culture of Koryaks. St. Petersburg: Nauka (in Russian).

Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC. 1997a. Y chromosome polymorphisms in Native American and Siberian populations: identification of founding Native American Y chromosome haplotypes. Hum Genet 100:536–543.

Lell JT, Schurr TG, Brown MD, Sukernik RI, Starikovskaya YB, Scozzari R, Cruciani F, Wallace DC. 1997b. Y-chromosome compound haplotype analysis: Implications for the peopling of the Americas and elucidation of intra-Siberian relationships. Am J Hum Genet 61(Suppl):A204.

Levin MG. 1958. Ethnic anthropology and problems of the ethnology of the peoples of the Far East. Moscow/Leningrad (in Russian).

Levin MG, Vasiliev BA. 1964. The Evens. In: Levin MG, Potapov LP, editors. The peoples of Siberia. Chicago: University of Chicago Press. p 670–684.

Lorenz JG, Smith DG. 1994. Distribution of the 9-bp mitochondrial DNA region V deletion among North American Indians. Hum Biol 66:777–788.

Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among Native North Americans. Am J Phys Anthropol 101:307–323.

Lum JK, Rickards O, Ching C, Cann RL. 1994. Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum Biol 66:567–590.

Malyarchuk BA, Derenko MV, Solovenchuk LL. 1994. Restriction pattern types of the major noncoding region of mitochondrial DNA in native and migrant inhabitants of northeast Asia. Russian J Genet 30:750–754.

Merriwether DA, Rothhammer F, Ferrell RE. 1994. Genetic variation in the New World: ancient teeth, bone, and tissue as sources of DNA. Experientia 50:592–601.

Merriwether DA, Rothhammer F, Ferrell RE. 1995. Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 98:411–430.

Merriwether DA, Hall WW, Vahlne A, Ferrell RE. 1996. mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet 59:204–212.

Mochanov YA. 1962. Early Neolithic of the Aldan River basin. Moscow: Nauka (in Russian).

Murashko O. 1994. A demographic history of the Kamchadal/Itel'men of the Kamchatka Peninsula: Modeling the precontact numbers and postcontact depopulation. Arctic Anthropol 31:16–30.

Nei M. 1972. Genetic distances between populations. Am Naturalist 106:283–292.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Nei M, Jin L. 1989. Variances of the average number of nucleotide substitutions within and between populations. Mol Evol Biol 6:290–300.

Nei M, Tajima F. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction site data. Genetics 105:207–217.

Passarino G, Semino O, Modiano G, Santachiara-Benerecetti AS. 1993. COII/tRNA^{Lys} intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (Southern Nepal) have oriental affinities. Am J Hum Genet 53:609–618.

Petrishchev VN, Kutueva AB, Rychkov YG. 1993. Deletion-insertion polymorphisms in ten mongoloid populations of Siberia: frequency of deletion correlates with the geographical coordinates of the area. Genetika 29:1196–1204 (in Russian).

Reynolds JB, Weir BS, Cockerham CC. 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. Genetics 105:767–779.

Ribiero-Dos-Santos AKC, Santos SEB, Machado AL, Guapindaia V, Zago MA. 1996. Heterogeneity of mitochondrial DNA haplotypes in pre-Columbian natives of the Amazon region. Am J Phys Anthropol 101: 29–37.

Rogers RA, Rogers LA, Hoffmann RS, Martin LD. 1991. Native American biological diversity and the biogeographic influence of Ice Age refugia. J Biogeography 18:623–630.

Saitou N, Imanishi T. 1989. Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. Mol Biol Evol 6:514–525.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Santos MR, Ward RH, Barrantes R. 1994. mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66:963–977.

Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, Wallace DC. 1990. Amerindian mitochondrial DNAs have rare Asian variants at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623.

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH. 1993. mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53:549–562.

Simchenko YB. 1976. Culture of reindeer herders in northern Eurasia. Moscow: Nauka (in Russian).

Skorik PY. 1965. Chukotko-Kamchatkan languages. Languages of northern USSR: Mongolian, Tungusic-Manchuric and Paleoasiatic languages. Leningrad: Nauka (in Russian).

Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. 1998. Mitochondrial DNA diversity in Chuchi and Siberian Eskimos: Implications for the genetic history of ancient Beringa and the peopling of the New World. Am J Hum Genet 63:1473–1491.

Stone AC, Stoneking M. 1994. Ancient DNA from a pre-Columbian Amerindian population. Am J Phys Anthropol 92:1158–1162.

Sukernik RI. 1992. Non-random distribution of GM haplotypes in northern Siberia. In: Roberts DF, Fujiki N, Torizuka K, editors. Isolation, migration and health. Cambridge: Cambridge University Press. p 107–111.

Sukernik RI, Lemza SV, Karaphet TM, Osipova LP. 1981. Reindeer Chukchi and Siberian Eskimos: studies on blood groups, serum proteins, and red cell enzymes with regard to genetic heterogeneity. Am J Phys Anthropol 55:121–128.

Sukernik RI, Osipova LP, Karaphet TM, Kirpichnikov GA. 1986. GM haplotypes and their frequencies in ten Chukchi populations. Genetic structure of Reindeer Chukchi. Genetika 22:2361–2368 (in Russian).

Sukernik RI, Schurr TG, Starikovskaya EB, Wallace DC. 1996. Mitochondrial DNA variation in native Siberians, with special reference to the evolutionary history of Americans Indians. Studies on restriction polymorphism. Genetika 32:432–439 (in Russian).

Swofford D. 1994. Phylogenetic Analysis Using Parsimony (PAUP), version 3.1.1. Champaign, IL: Illinois Natural History Survey.

Szathmáry EJE. 1981. Genetic markers in Siberia and northern North American populations. Yrbk Phys Anthropol 24:37–73.

Torroni A, Schurr TG, Yang C-C, Szathmary EJE, Williams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM, Wallace DC. 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the Na-Dené populations were founded by two independent migrations. Genetics 130:153–162.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. 1993a. Asian affinities and the continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590.

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MA, Crawford MH, Comuzzie AG, Wallace DC. 1993b. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53:591–608.

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC. 1994a. A mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci U S A 91:1158–1162.

Torroni A, Chen Y-S, Semino O, Santachiara-Benerecetti AS, Scott CR, Lott MT, Winter M, Wallace DC. 1994b. mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54:303–318.

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang JG, Droma T, Wallace DC. 1994c. Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93:189–199.

Torroni A, Lott MT, Cabell MF, Chen Y-S, Lavergne L, Wallace DC. 1994d. mtDNA and the origin of Caucasians: Identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. Am J Hum Gen 55:760–776.

Torroni A., Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu d, Savontaus M-L, Wallace DC. 1996. Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850.

Vasilievskiy RS. 1971. The origins and ancient culture of the Koryak. Novosibirsk: Nauka (in Russian).

Vdovin IS. 1973. Essays on the ethnohistory of the Koryaks. Moscow: Nauka (in Russian).

Wallace DC, Garrison K, Knowler WC. 1985. Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol 68:149–155.

Ward RH, Frazier BL, Dew-Jager K, Pääbo S. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci U S A 88:8720–8724.

Ward RH, Redd A, Valencia D, Frazier B, Pääbo S. 1993. Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci U S A 90:10663–10667.